TRADING SYSTEMS

AND

METHODS

+Website

FIFTH EDITION

PERRY J. KAUFMAN

Trading Systems and Methods

Founded in 1807, John Wiley & Sons is the oldest independent publishing company in the United States. With offices in North America, Europe, Australia and Asia, Wiley is globally committed to developing and marketing print and electronic products and services for our customers' professional and personal knowledge and understanding.

The Wiley Finance series contains books written specifically for finance and investment professionals as well as sophisticated individual investors and their financial advisors. Book topics range from portfolio management to e-commerce, risk management, financial engineering, valuation and financial instrument analysis, as well as much more. For a list of available titles, visit our Web site at www.WileyFinance.com.

Trading Systems and Methods

Fifth Edition

PERRY J. KAUFMAN

WILEY

John Wiley & Sons, Inc.

To Barbara

Contents

Preface to the Fifth Edition

In the past eight years, since the last edition, our industry has continued to change. The extraordinary bull market of the late 1990s, followed by the bursting of the tech bubble in 2000 seemed to be events that could never be overshadowed, but the subprime collapse in 2007 proved us wrong. We learned what risk was all about, when money was pulled from every possible investment at the same time. In many cases, the investments that were liquidated had no other relationship than having profits that were needed to cover losses elsewhere. The principles of diversification held true, but we saw the worst-case scenario, where everything moved in the same way at the same time. It was an event with a very low probability, but not zero.

During the years that have followed, we would expect much more focus on risk management, rather than risk measurement. Understanding how to reduce risk before the fact is much more productive than identifying it afterward. Some hedge funds, following in the steps of Long-Term Capital Management, have chosen to see this as a rare event, not likely to be repeated. The rationale for this is that, in order to reduce the chances of large risk, you must also reduce returns. They see investors as preferring the small chance of a large loss to the less acceptable assurance of lower profits. I won't try to judge the merits of this decision.

On the other hand, we should all understand the best choices for controlling risk. With that in mind, many of the changes in this edition address risk control, from the individual trade level, to the strategy rules, to the portfolio.

COHERENCE

One of the improvements in this edition is the added coherence from one section to another and from one chapter to another. There will be references, both forward and backward, showing similarities between many techniques. By incorporating those references, some of the duplication has been removed. Considerable effort was made to use the same notation throughout the book, in hope that it will make the formulas easier to understand. You will also find that there is a greater attempt to make this material flow from section to section as a continuous learning process.

MORE STRATEGIES, MORE PROGRAMS AND SPREADSHEETS

Each year brings new ideas, and many articles and books have shed light on new techniques or better ways to approach an old problem. Wherever possible, those ideas have been added here, with references to the original material. There is more cross-pollination with the securities industry, and you will find more terms now used by both futures and stock traders.

Examples make learning easier, and this edition has many more examples, along with more programs and spreadsheets that will help you take whichever ideas are appealing and try them on your own. Wherever possible, the spreadsheet code uses "offset" to allow the calculation periods to be changed. These examples continue to use TradeStation and Excel, which remain the most popular tools. While there are many other choices, code from these two sources can be easily converted to other programs.

UPDATED CHARTS

Along with more examples, many of the old examples and charts have been brought up to date. While there may be an historic interest in market patterns during the 1970s and 1980s, the recent 10 years provide dramatic price movements and seem more relevant. We would all agree that it's good for a strategy to have been profitable in the 1980s, but more important that it succeeded during the past 5 or 10 years. In many cases, the old patterns are still unique and should not be ignored, but every book has its limits.

SEARCHING FOR ROBUSTNESS

The goal of a system developer and/or a trader is to find or create a trading method that will work in many different situations, hopefully across many different markets, and keep working for as long as possible. A solution that is robust satisfies those objectives. Because if its importance, there are comments throughout the book addressing the robustness of various methods and ways to enhance that quality. Chapter 21, System Testing, addresses this directly, but it is not the only place.

Robustness is an easy concept to understand, but a robust strategy has a return and risk profile that is not as attractive as one that has been fitted to the data. Success with fewer rules over more markets and data yields robustness, but at the price of lower returns and higher risk. It is necessary to understand and embrace the natural risk of a strategy and market in order to succeed in the long term. If you try to engineer all of the risk out of a trade, it will only surface somewhere else when it is least welcome.

COMPANION WEBSITE

The companion website has been greatly expanded with both TradeStation programs (Version 9) and Excel 2010 spreadsheets. It is expected that MetaStock code will be added in the near future. There is also a list of contents at the back of this book. Whenever a website program relates to a section of the book, there is an icon in the margin to remind you of its availability. Hopefully, this will make the development and verification of new ideas more convenient.

WITH APPRECIATION

This book draws on the hard work and creativity of hundreds of traders, financial specialists, engineers, and many others who are passionate about the markets. They continue to redefine the state-of-the-art and provide all of us with both profitable techniques and valuable tools.

A long overdue thanks to Janette Perez of TradeStation for her generous help. My gratitude to Pamela van Giessen and Emilie Herman of John Wiley & Sons, who continue to provide immeasurable help and encouragement. And to my wife, Barbara, whose everlasting support is only enhanced by rolling her eyes whenever I say that this is my last book, ever.

As a final note, I would like to thank all the previous readers who sent messages about typographical errors, omissions, and just simple errors. They have all been corrected. It makes this edition that much better.

Perry J. Kaufman
Freeport, Grand Bahama
November 2012

Introduction

*It is not the strongest of the species that survive, nor the most intelligent,
but the ones most responsive to change.*

—Charles Darwin

Let's start by redefining the term *technical analysis*. Technical analysis is the systematic evaluation of price, volume, breadth, and open interest, for the purpose of price forecasting. A systematic approach may simply use a bar chart and a ruler, or it may use all the calculation power available. Technical analysis may include any quantitative analysis as well as all forms of pattern recognition. Its objective is to decide, in advance, where prices will go over some time period, whether 1 hour, 1 day, or 5 years. Technical analysis must have clear and complete rules.

Technical analysis is no longer just the study of chart patterns or the identification of trends. It encompasses intramarket analysis, complex indicators, mean reversion, and the evaluation of test results. It can use a simple moving average or a neural network to forecast price moves. This book serves as a reference guide for all of these techniques, puts them in some order, and explains the functional similarities and differences for the purpose of trading. It includes some aspects of portfolio construction and multilevel risk control, which are integral parts of successful trading.

THE EXPANDING ROLE OF TECHNICAL ANALYSIS

Quantitative methods for evaluating price movement and making trading decisions have become a dominant part of market analysis. Those who do not use methods such as overbought and oversold indicators are most likely to watch them along the bottom of their screen. The major financial networks are always pointing out price trends and double bottoms, and are quick to say that a price move up or down was done on low volume to show

1

that it might be unreliable. The 200-day moving average seems to be the benchmark for trend direction. These comments show the simplicity and the acceptance of technical analysis.

Events beginning in 2002 cast doubt on the integrity of the research produced by major financial houses that have a conflict between financing/underwriting and retail brokerage. The collapse of Enron has caused us to question the earnings, debt, quality-of-business, and other company data released to the public by large and small firms. It is not surprising that more quantitative trading methods have been adopted by research firms. When decisions are made with clear rules and calculations that can be audited, those analysts recommending buys and sells are safe from scrutiny.

Extensive quantitative trading exists around the world. *Interest rate arbitrage* is a major source of revenue for banks. *Location arbitrage* is the process that keeps the price of gold and other precious metals the same all over the globe. *Program trading* keeps the price of the overall stock market from diverging from S&P futures and SPY (the SPDR ETF) prices. Recently these fully automated systems have been called *algorithmic trading*.

If you don't think of arbitrage as technical trading, then consider market neutral strategies, where long and short positions are taken in related markets (pairs trading) in order to profit from one stock rising or falling faster than the other. If you change your time horizon from hours and days to milliseconds, you have high frequency trading. You might prefer to take advantage of the seasonality in the airline industry or try your hand trading soybeans. Both have clear seasonal patterns as well as years when other factors (such as a disruption in energy supply) overwhelm the seasonal factors. Trading seasonal patterns falls under technical analysis.

Technology that allows you to scan and sort thousands of stocks, looking for key attributes—such as high momentum, a recent breakout, or other indicator values—is also technical analysis on a broader scale. High frequency trading, arbitrage that lasts only milliseconds, has become a profit center for large financial institutions, but involves placing computer equipment as close to the source of the exchange price transmission as possible—a contentious issue. High frequency trading is credited for adding liquidity by increasing volume in equities trading, but has also been blamed (perhaps unfairly) for spectacular, highly volatility price moves.

Most impressive is the increase in managed funds that use technical and quantitative analysis. Many billions of investment dollars are traded using trend-following systems, short-term timing, mean reversion, and countless other techniques. It is thought that well over half of all managed money uses algorithmic trading. Technical analysis allows you to backtest and estimate the expected risk, two great advantages to the fund manager. The use of technical analysis has infiltrated even the most guarded fundamental fortresses.

CONVERGENCE OF TRADING STYLES IN STOCKS AND FUTURES

The development of technical analysis has taken a different path for stocks and futures. This seems natural because the two markets cater to investors with different time frames

and different commercial interests. At the same time, the markets place very different financial demands on the investor.

The original users of the futures markets were grain elevators and grain processors, representing the supply side and the demand side. The elevators are the grain whole-salers who bought from the farmers and sold to the processors. The futures markets represented the fair price, and grain elevators sold their inventory on the Chicago Board of Trade in order to lock in a price (hopefully a profit). The processors, typically bread manufacturers or meat packers, used the futures markets to lock in a low price for their material cost and as a substitute for holding inventory. Both producer (the sell side) and processor (the buy side) only planned to hold the position for a few weeks or a few months, until they either delivered their product to market or purchased physical inventory for production. There was no long-term investment, simply a hedge against risk. Futures contracts, similar to stock options, expire every two or three months and can be held for about one year; therefore, it is nearly impossible to "invest" in futures.

One other critical difference between futures and stocks is the leverage available in futures. When a processor buys one contract of wheat, that processor puts up a good faith deposit of about 5% of the value of the contract. If wheat is selling for $10.00 a bushel and a standard contract is for 5,000 bushels, the contract value is $50,000. The processor need only deposit $2500 with the broker. The processor is essentially buying with leverage of 20:1.

In the 1970s, the futures trader paid an outrageous round-turn commission of $50 per contract. This is about 0.3 of 1 percent, far less than the stock market cost of 1%, but one of the highest commission ratios in the futures industry. Now, years after negotiated commissions have become part of the system, the fee is closer to $8, or 0.05 of one percent. Commission costs are so low that they are not a consideration when trading futures. To be fair, the cost of trading equities has also dropped in the same way, but favors those trading larger positions.

How do the high leverage and low commissions affect trading in futures? Low costs allow short holding periods. Floor traders don't invest—they look to scalp the market or capture a fast, volatile price move. In the derivatives markets, *fast* is 1 to 3 days, and *slow* is anything longer than 30 days.

Although speculation has always had a place in the stock market, the investor, rather than the trader, has been the major force. The stock market is an investment in America. The growth of the economy parallels the growth and efficiency of industry. Of course, commissions and tax regulations played a large part in shaping the long-term view of the investor. When commission costs were 1% for each buy and sell order, it was not possible to be a short-term trader. That role was reserved for the market maker on the floor of the stock exchange. It is difficult to be a trader of any sort when you pare 2% from each of your trades. Even now, some mutual funds charge high fees or penalties for liquidating a position before six months or one year. In addition, favorable tax treatment strongly encouraged holding positions for at least six months to satisfy the long-term capital gain rule. The uptick rule for selling discouraged speculating on the decline of stock prices, and while it is not in effect at this time, politicians are inclined to reinstate it in the belief

that it will reduce market volatility. Even now, short sales are not allowed in most retirement funds. To get around these rules, exchange-traded funds (ETFs) such as SPY and QQQ allow buying and short selling with no expiration date and low cost. The main differences between trading ETFs and futures is that futures allow leverage, trade in larger size, expire at fixed intervals, and are guaranteed by a major institution, such as the Chicago Mercantile Exchange. Because of the low cost in stocks and the new trading vehicles in the form of ETFs, stock traders now look to the methods used by futures traders to identify trends and mean-reverting opportunities faster and use tighter risk controls.

A LINE IN THE SAND BETWEEN FUNDAMENTALS AND TECHNICAL ANALYSIS

The market is driven by fundamentals. These are often employment, GDP, inflation, consumer confidence, supply and demand, or geopolitical factors—all of which create expectations of price movement. But it is just too difficult to trade using those facts, and economists have never been very accurate. Economic reports are not usually timely, and individual companies are not forthcoming about problems. We have had too many cases where the data we use to make fundamental decisions about individual companies have been unreliable. We can add that to the conflict of interest inherent in the government's calculation of the Consumer Price Index, because an increase in the CPI requires that all those receiving Social Security checks get a cost-of-living increase.

Technical analysis, when used to determine the long-term direction of prices, attempts to objectively evaluate these complex fundamentals. It is no different from the economists who use regression, seasonal, and cyclic analysis to forecast the economy. The technical trader can use those tools as well as chart trendlines, pattern recognition, and probability distributions. Perhaps the economists are doing the same thing.

It is well known that the Federal Reserve monitors trading and prices in order to decide how to time their interest rate changes and, when necessary, their currency intervention. All monetary authorities know that when their currency is rising too fast, you don't try to stop it. If the public wants to buy the Japanese yen, the Central Bank doesn't have enough clout to stop it unless it first waits for the move to be exhausted. It must use its resources carefully, and it uses market know-how and price analysis to time its actions.

The primary advantages of a technical approach are that it is objective and completely well-defined. The accuracy of the data is certain. One of the first great advocates of price analysis, Charles Dow, said:

> *The market reflects all the jobber knows about the condition of the textile trade; all the banker knows about the money market; all that the best-informed president knows of his own business, together with his knowledge of all other businesses; it sees the general condition of transportation in a way that the president of no single railroad can ever see; it is better informed on crops than the farmer or even*

> *the Department of Agriculture. In fact, the market reduces to a bloodless verdict all knowledge bearing on finance, both domestic and foreign.*

Much of the price movement reflected in any market is anticipatory; it results from the expectations of the effects of macroeconomic developments or the outcome of good corporate management and new products. Markets, however, are subject to change without notice. For example, the government may block the merger of two companies, or approve or reject a new drug. A hurricane bound for the Philippines will send sugar prices higher, but if the storm turns off course, prices reverse. Anticipation of employment reports, housing starts, or corn production reports causes highly publicized professional estimates, which may correctly or incorrectly move prices before the actual report is released. Markets then react to the accuracy of the estimates rather than to the economic data itself. By the time the public is ready to act, the news is already reflected in the price.

PROFESSIONAL AND AMATEUR

Beginning technical traders may find a system or technique that seems extremely simple and convenient to follow, one that appears to have been overlooked by the professionals. Most often there is a simple reason why that method is not used. As you learn more about trading, you find that execution is difficult, or the risk is much higher than originally expected, or that the system has too many losses in a row. Trading is a business, not one to be taken casually. As Richard Wyckoff said, "Most men make money in their own business and lose it in some other fellow's." Plan to invest your time before your money, so that when you begin trading, you have more realistic expectations.

That does not mean that simple systems don't work, but that each has a return and risk profile that is typical of that style and difficult to change. One purpose of this book is to present many different trading methods, each with its own risk and reward profile, so that each trader understands the true cost of trading.

To compete with a professional speculator you must be more accurate in anticipating the next move. This can be done by

- Recognizing recurring patterns in price movement and determining the most likely results of such patterns.
- Identifying the "trend" of the market by isolating the basic direction of prices over a selected time interval.

The bar chart, discussed in Chapter 3, is the simplest representation of the market. These patterns are the same as those recognized by Jesse Livermore, in the early 1900s, on the ticker tape. Because they are interpretive, more precise methods such as point-and-figure charting are also used, which add a level of exactness to charting. Point-and-figure charts are popular because they offer specific trading rules and show formations similar to both bar charting and ticker-tape trading.

Mathematical modeling, using traditional regression or statistical analysis, has become a popular technique for anticipating price direction. Most modeling methods are modifications of developments in econometrics and basic probability and statistical theory. They are precise because they are based entirely on numerical data; however, they need trading rules to make them operational.

The proper assessment of the price trend is critical to most trading systems. Countertrend trading, which takes a position opposite to the trend direction, is just as dependent on knowing the trend as a trend-following technique. Large sections of this book are devoted to the various ways to identify the trend, although it would be an injustice to leave the reader with the idea that a "price trend" is a universally accepted concept. There have been many studies published claiming that price trends do not exist. The most authoritative papers on this topic are collected in Cootner, *The Random Character of Stock Market Prices*; very readable discussions can be found in the *Financial Analysts Journal*, an excellent resource.

Personal money management has an enormous number of tools, many of which can be found in Excel and other spreadsheet software. These include linear regression and correlation analysis. There is also inexpensive software to perform spectral analysis and apply advanced statistical techniques. There is an Excel add-in, *Solver*, that can easily be adapted to portfolio allocation. Trading systems development software, such as TradeStation and MetaStock, have provided platforms and greatly reduced the effort needed to program your ideas. Professionals maintain the advantage of having all of their time to concentrate on the investment problems; however, nonprofessionals are no longer at a disadvantage.

RANDOM WALK

It has been the position of many advocates of fundamental and economic analysis that there is no sequential correlation in the direction of price movement from one day to the next; that is, prices have no memory of what came before—this has been named the *random walk* theory. Prices will seek a level that will balance the supply-demand factors, but that level will be reached either instantaneously, or in an unpredictable manner, as prices move in response to the latest available information or news release.

If the random walk theory is correct, the many well-defined trading methods based on mathematics and pattern recognition will fail. The problem is not a simple one, but one that should be resolved by each system developer because it will influence the type of systematic approaches that best suits them. There are two arguments against random movement in prices.

The first argument is simply the success of many algorithmic trading strategies. There is definitive documentation of performance for systematized arbitrage programs, hedge funds, and derivatives funds, showing success for upwards of 20 or 30 years. This is not to say that all technical programs are successful—far from it. But neither

are fundamental methods. You still need a sound strategy, whether discretionary or systematic, in order to be profitable. Not everyone can create and implement such a strategy.

The second argument against the random walk is that prices move on anticipation. One can argue academically that all participants (the "market") know exactly where prices should move following the release of news. However practical or unlikely this is, it is not as important as market movement based on anticipation of further movement. For example, if the Fed lowered rates twice this year and the economy has not yet responded, would you expect it to lower rates again? Of course you would. Therefore, as soon as the Fed announces a rate cut you would speculate on the next rate cut. When most traders hold the same expectations, prices move quickly to that level. Prices then react to further news relative to expectations. Is this price movement that conforms to the random walk theory? No. But the actual pattern of price movement can appear similar to random movement.

Excluding anticipation, the apparent random movement of prices is dependent on both the time interval and the frequency of data observed. When a long time span is used, from 1 to 20 years, and the data averaged to enhance the smoothing process, the trending characteristics appear more clearly, along with seasonal and cyclic variations. Technical methods, such as moving averages, are often used to isolate these price characteristics. Averaging daily or weekly data to create monthly or quarterly prices smooth out irregular short-term movements, resulting in higher correlations between successive prices. With less frequent data, it is easier to see a trend. In general, the use of daily data shows more noise (random movement) than weekly or monthly data.

In the long run, prices seek a level of equilibrium. For stocks, equilibrium is where the return on investment (appreciation of share value plus dividends), balanced with the risk of the investment, puts it on an equal footing with the returns of a risk-free investment, such as Treasury notes. In futures, equilibrium is the balance between supply and demand.

Prices do not move in a symmetric pattern, and they do not have a normal distribution: two additional facts that argue against random walk. The asymmetry of the index markets, in particular those built on traditional stocks, is easy to understand because the public consists overwhelmingly of buyers. But it is also the nature of price movement to show unique patterns when prices move farther from their normal value during periods of stress, or exceptional supply and demand imbalance. When looking at price movement in terms of "runs"—hours or days when prices continue in the same direction for an unusually long sequence—we find that price data has a *fat tail*, representing much longer runs than can be explained by a normal distribution. The existence of a fat tail also means that some other part of the distribution must differ from the norm because the extra data in the tail must come from somewhere else. Throughout this book, we refer to these differences in price patterns as the reason why certain trading methods work.

Price movement is driven by people, and people can buy and sell for nonrandom reasons, even when viewed in large numbers. For example, an investment fund will enter the

market without regard to timing, based on their monthly additions or redemptions. This in turn moves prices and creates opportunities that allow traders to profit. The long-term trends that reflect economic policy, normally identified by quarterly data, can be of great interest to longer-term *position traders.* It is the short-term price movements caused by anticipation (rather than actual events), extreme volatility, prices that are seen as far from value, countertrend systems that rely on prices reversing direction, and those that attempt to capture trends of less duration that are the primary focus of this book.

DECIDING ON A TRADING STYLE

It may seem backwards to talk about a trading style in advance of reading all the material, but many traders have already decided that they want to day trade or hold long-term positions because it suits their disposition, their belief of what moves prices, or their time schedule. With that in mind, short-term and long-term traders will focus on different strategies and markets, while portfolio structure and risk control will be much the same for either approach.

To understand how markets and different trading styles work together, consider a daily chart of any market, an individual stock, a short-term interest rate futures contract, or the S&P 500 index. There are periods of trending and sideways patterns. However, if you change that chart from daily to weekly, or daily to monthly, the long-term trend emerges. It is much easier to see the trend when you use fewer data points over a longer time interval. The presentation of the chart smoothes the appearance of the data (see Figures 1.1 and 1.2).

Now go in the other direction, using 20-minute bars instead of a daily chart. The trend is more difficult to see. What appeared to be a smooth period leading up to the peak
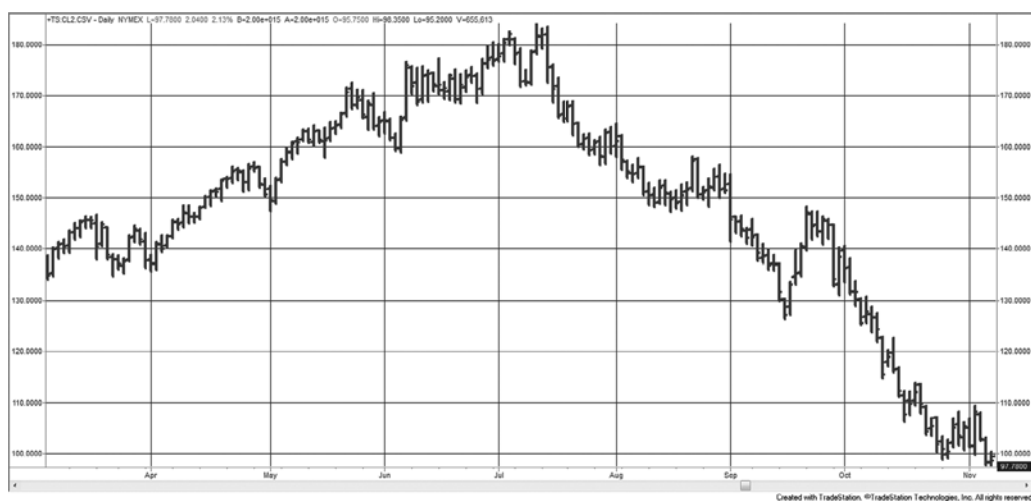
FIGURE 1.1 Crude oil daily chart with July 2008 in the center.
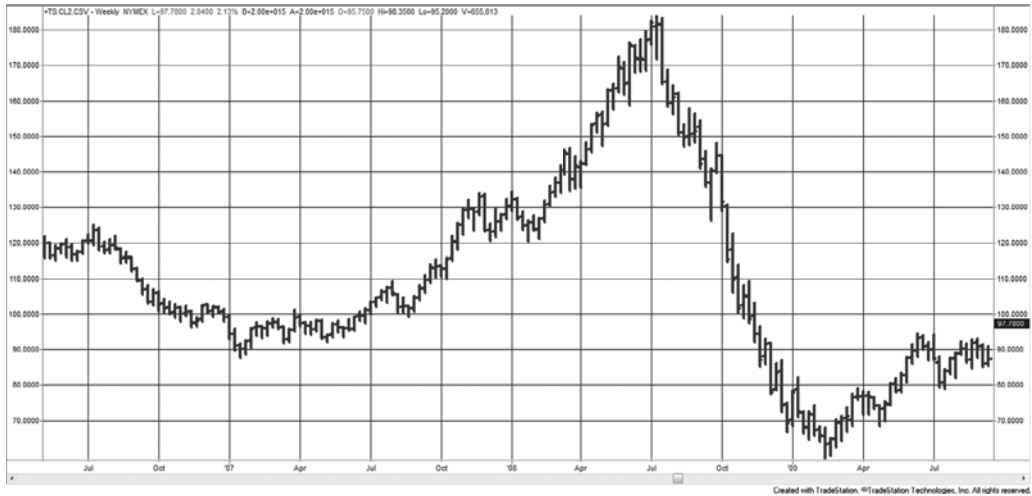
FIGURE 1.2 Crude oil weekly chart with July 2008 to the right of center.

in July 2008 (Figure 1.3), now looks very erratic. As the bars become shorter, the price noise appears to increase.

Selecting a price frequency that complements your trading strategy is often ignored by traders. If you are a long-term, macrotrend follower, then you want the price series that shows more trends, which are improved by monthly, weekly, or daily charts, although monthly is generally too low frequency for traders. Short-term traders focus on mean reversion or fast directional price moves, and that strategy is enhanced using higher-frequency data, such as hourly or 15-minute bars.

FIGURE 1.3 Crude oil 20-minute chart with July 2008 in the center.

MEASURING NOISE

The need to select the data frequency that best suits the strategy can be verified by measuring price *noise*. Noise is the erratic movement that surrounds the underlying direction of prices at any time. High noise can be compared to a drunken sailor's walk, while low noise is a straight line from the starting to the ending point.

There are a number of ways to measure noise, including *price density, efficiency ratio* (also called *fractal efficiency*), and *fractal dimension*. It is important that these measurements remove volatility because noise should not be confused with volatility. In Figure 1.4 a short, hypothetical period of price movement gives an example of noise measured by the efficiency ratio (ER). ER is calculated by dividing the net move (the change for point A to point B) by the sum of the individual moves during that period, each taken as positive numbers.

$$Efficiency\ ratio = \frac{Net\ change\ in\ price\ (as\ a\ positive\ number)}{Sum\ of\ individual\ price\ changes\ (as\ positive\ numbers)}$$

or

$$ER_t = \frac{|P_t - P_{t-n}|}{\sum_{t=t-n}^{i=t}|P_i - P_{i-1}|}$$

where n is the calculation period.

Figure 1.5 illustrates the relative level of noise that might comprise a price move of the same net change. The straight line indicates no noise, the smaller changes that move above and below the straight line would be medium noise, and the large swing are high noise. However, in this example it is not possible to distinguish the level of noise from volatility, yet they are not the same. In Figure 1.6, the net change in price is from 440 to
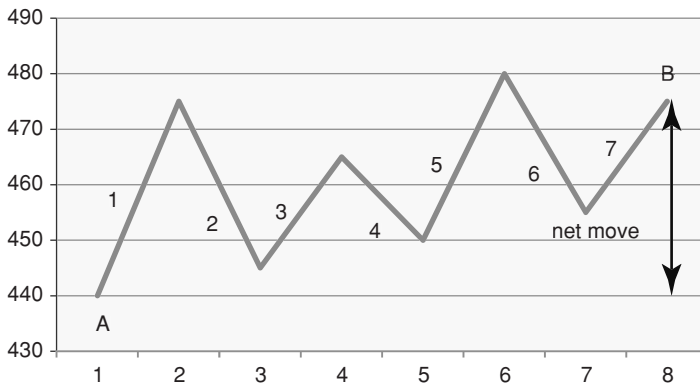
FIGURE 1.4   Basic measurement of noise using the efficiency ratio (also called fractal efficiency).

Levels of Noise
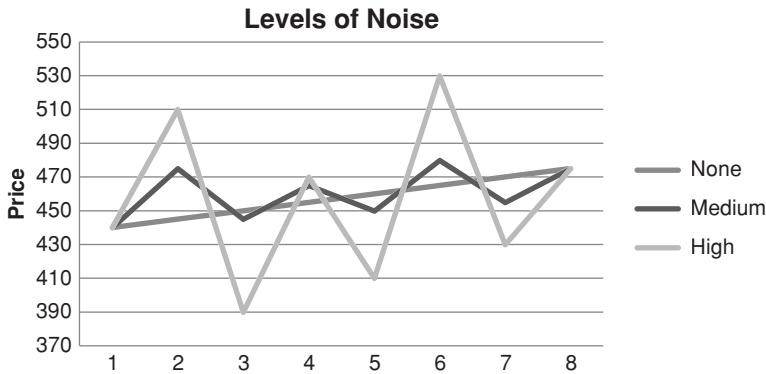
FIGURE 1.5   Three different price patterns all begin and end at the same point. The straight line shows no noise, the smaller variations are medium noise, and the larger swings are high noise.

475 in one case and from 440 to 750 in the other, yet the sum of the individual component changes are similar, 595 and 554. The efficiency ratio is 0.06 for the first and 0.56 for the second, showing that the first is very noisy while the second has relatively low noise (see Table 1.1). Remember that a ratio near 1 shows a strong trend, and a ratio near 0 only noise. If prices are moving up quickly, then even large swings may not be considered a serious interference with the trend.

Other Measurements of Noise

The previous example of noise used the efficiency ratio; however, price density and fractal dimension may also be used. Intuitively, *price density* can be seen as the extent to

Same Volatility, more gain = less noise
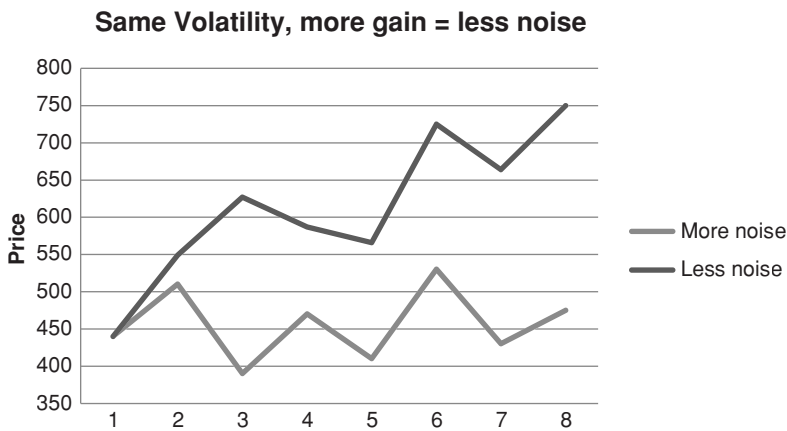
FIGURE 1.6   By changing the net price move, we can distinguish between noise and volatility. If the sum of the individual price changes are the same, but the net move is larger, then the noise is less.

TABLE 1.1 These price changes, reflecting the patterns in Figure 1.6, show that larger individual price changes do not correspond to higher noise if the net change over the entire period is much larger.

Day	High noise	Low noise	Diff high	Diff low
1	440	440		
2	510	549	70	109
3	390	627	120	78
4	470	587	80	40
5	410	566	60	21
6	530	725	120	159
7	430	664	100	61
8	475	750	45	86
Net change	35	310	595	554
Noise			0.06	0.56

which prices fill a box. If we take a 10-day period of price movement charted with highs and lows, and draw a box touching the highest high and lowest low, then the density is how much of that box is filled. This is measured as

$$Price\ Density = \frac{\sum_{i=t-n+1}^{i=t}(High_i - Low_i)}{Max\,(n\text{-}day\ high) - Min\,(n\text{-}day\ low)}$$

Fractal dimension cannot be measured exactly but can be estimated over n days using the following steps:

1. Max = highest high over n days
2. Min = lowest low over n days
3. Range = max – min
4. $dx^2 = \left(\dfrac{1}{n}\right)^2$
5. $L = \sum_{i=t-n+1}^{i=t}\sqrt{dx^2 + \dfrac{p_i - p_{i-1}}{Range}}$
6. $FD = 1 + \dfrac{\ln(L) + \ln(2)}{\ln(2 \times n)}$

There is a strong relationship between fractal dimension and the efficiency ratio (fractal efficiency), and there is a similarity in the construction of price density and fractal dimension. In step 5 the term inside the square root sign accumulates the change in price relative to the range over the calculation period. Of the three methods of measuring noise, the efficiency ratio seems to be the clearest, and that will be used in the following analyses.

Impact on Trading

To determine the significance of the efficiency ratio, a 20-day average noise was calculated for a wide range of futures and world equity index markets for January 2000 through March 2012. A corresponding 40-day moving average trend-following strategy was applied to the same markets (a complete discussion of trend systems can be found in Chapter 8). The trend system used the most basic rules, going long when the trendline turned up and short when it turned down. It was always in the market and there were no costs applied. The results of both the noise and corresponding trend results are scattered in Figure 1.7. Trend results are shown as a *profit factor*, gross profits divided by gross losses. Higher factors relate to better risk-adjusted returns. A simple regression line was drawn through the scatter diagram to emphasize the relationship.

Figure 1.7 shows a pattern from the bottom left to the top right of the chart. Profit factors under 1.0 are net losses, those above 1.0 are gains. The noise is greatest on the left (at 0.204) and lowest on the right (about 0.266). The results can be interpreted as follows: Low noise is good for trend following and high noise is not. That interpretation can be taken further as high noise favors mean-reverting strategies. Tests over different time periods, such as the 1990s, will show much stronger trend than recently, and can shift returns higher, but the relationship of noise to success will remain the same.

A closer look at the results shows that the markets in the top right are short-term interest rates, which are closely tied to Central Bank policy. The next trendiest markets
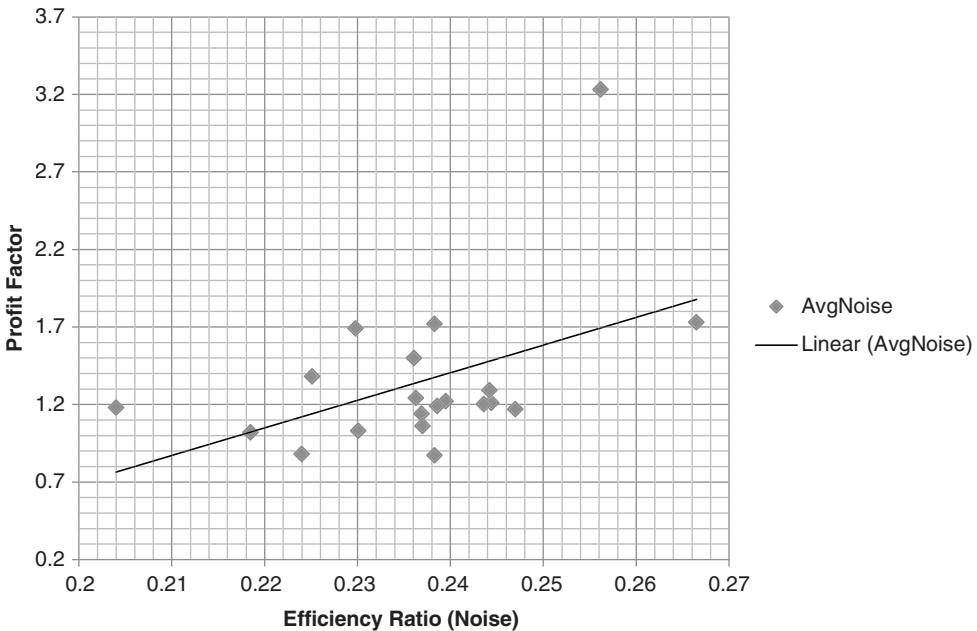
FIGURE 1.7 A scatter of the average 40-day noise and the information ratio based on a simple 40-day trend system, over 12 years from 2000.

are the longer maturity rates, then USD crossrates, energy markets, and metals. At the far lower left on the chart are the equity index markets. Equity index markets have the greatest noise and the worse trend-following performance of all sectors.

This concept will be extremely important when deciding on a trading style. Long-term traders, those interested in macrotrends, should combine low frequency data and long-term trends. Short-term traders should use high frequency data and favor mean-reverting strategies. Of course there are many exceptions to this approach, and opportunities are everywhere. This relationship of price noise, market patterns, and data frequency is intended to put a framework around the most common decisions made when developing trading strategies.

MATURING MARKETS AND GLOBALIZATION

The level of noise in each market can tell us about the maturity of that market and the nature of traders actively using it. The U.S. equity markets are where companies go to finance their operations. Typical U.S. workers participate in the equity markets indirectly through their retirement programs, and many are actively involved in making the decisions in the allocation of those funds. The most conservative choose money markets or guaranteed government debt obligations; others allocate a portion to the overall market using S&P ETFs, and still others choose specific sectors or even individual stocks to allocate part of their resources.

Other countries are not as involved in their equity markets, even though movement of the equity index in these countries reflects the health of their economy. With less involvement, there is less liquidity and participation is limited to a narrow, less homogeneous group of investors or traders. However, most world markets are becoming more active, even if that liquidity comes from globalization, where traders from one country buy and sell shares in another country. If we look at the history of price noise as reflected in the North American equity index markets, there is a steady increase in noise over the past 20 years (see Figure 1.8). This corresponds to the increase in volume of those markets, reflected in the S&P cash index, SPX, shown in Figure 1.9.

This increase in volume, corresponding to the maturity of the market, is not restricted to the United States or Europe, but a general phenomenon affecting all markets. Figure 1.10 shows the pattern in 5-year intervals grouped by geographic region: greater Asia, Australia and New Zealand, Eastern Europe, Europe (including the U.K.), Latin America, North America, and South America. The left scale is the average noise of all markets in that group. The overall picture is that the markets in all regions are maturing and that this maturity can come swiftly, as shown in Eastern Europe. There is an interesting exception in Latin America (represented only by Mexico) where the noise value has increased, representing less liquidity. This would not happen if the group was diversified and is the result of decreasing confidence in the economy, hence less liquidity. The European region shows a move up in the past five years. Unlike in the United States, European
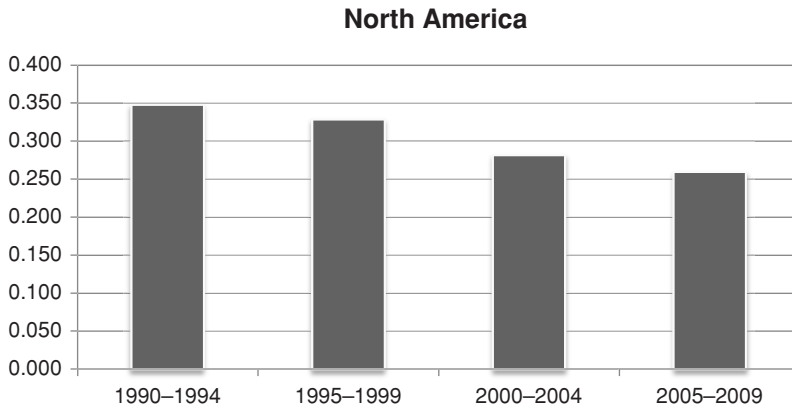
North America

FIGURE 1.8   By measuring the market noise of North American equity index markets during 5-year intervals, we can see the increase in noise over the past 20 years.
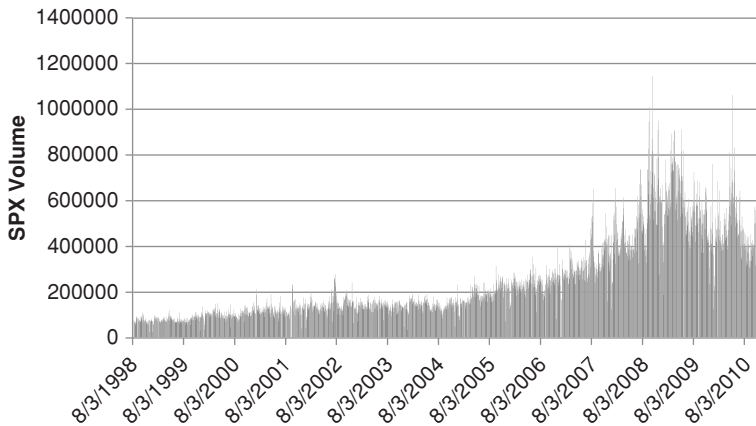
FIGURE 1.9   Volume traded of the S&P 500 index components for the most recent 10 years.

Market Usage by Region
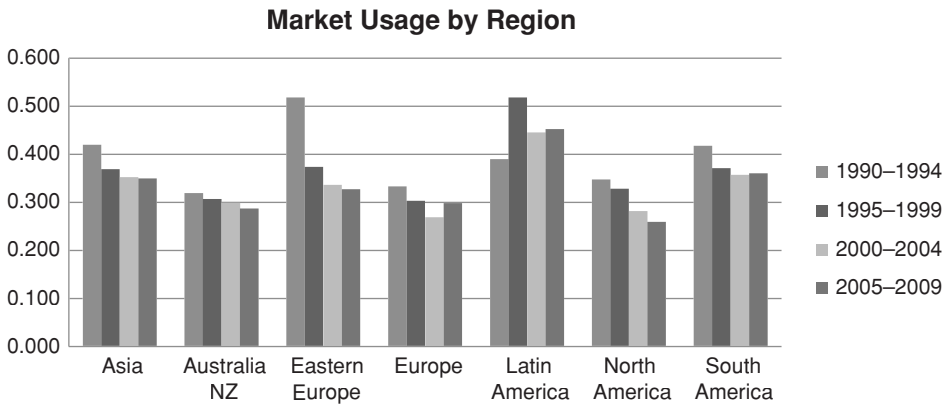
FIGURE 1.10   Relative change in maturity of world markets by region.
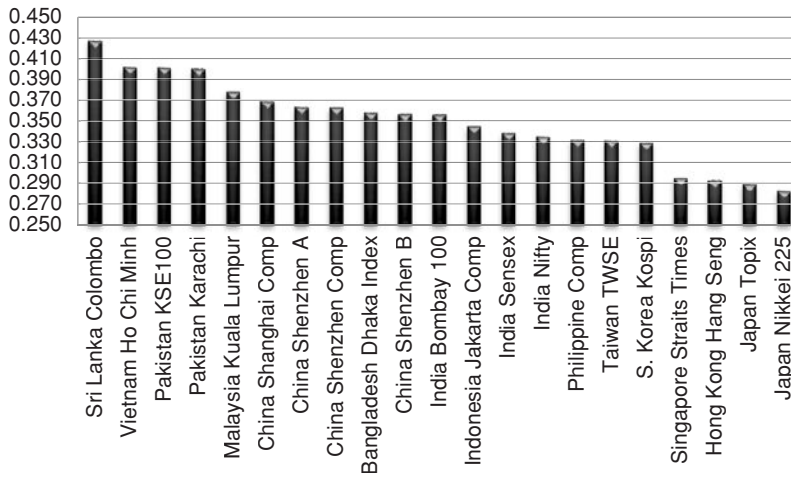
FIGURE 1.11 Ranking of Asian Equity Index Markets, 2005–2010.

investors are not as committed to the equity markets, and the economic crisis of the past few years has resulted in more investors exiting those markets than in the United States.

Asia is the most important area of world development at this time. China, which holds most of the U.S. debt, has given a great deal of economic freedom to its people, but limited access to the equity markets for outside investors. Figure 1.11, which is ranked from higher to lower noise values (less maturity to more maturity) from left to right, shows the relative development of the Asian equity markets. It is not surprising that Japan is the most developed, followed by Hong Kong, Singapore, South Korea, and Taiwan. These represent the most open economies in Asia. At the other end are Sri Lanka, Vietnam, Pakistan, and Malaysia, countries without access to global investors. India's Sensex shows greater participation than the China Shanghai Composite, but both are toward the center of the ranking. Those countries that allow greater access to traders in the futures will move toward the right in the ranking. For now, the Shanghai Composite shows an average value of about 0.37, while the North American markets are at 0.25.

BACKGROUND MATERIAL

The contents of this book assume an understanding of the stock market and futures markets, such as the S&P 500 and Treasury notes. These futures markets have a great impact on stock patterns and trade 24 hours a day. The rules and mechanics of those markets are not explained here unless they directly relate to a trading strategy. Ideally the reader should have read one or more of the available trading guides and should understand the workings of a buy or sell order and the contract specifications of futures. Experience in actual trading would be helpful. A professional trader, a broker, or a purchasing agent will already possess all the qualifications necessary, as will any businessperson

who understands how prices reflect earnings and the need to accumulate inventory at the lowest price. Individuals who manage their own stock portfolio or watch one of the financial news networks are also qualified. It also helps if you enjoy playing competitive games and you like to win.

There are excellent books available to both the beginning and advanced trader. The ones that stand out as valuable sources of general information are Jack Schwager's two-volume set, *Schwager on Futures*, which includes one volume on fundamental analysis and one on technical analysis. John Murphy's *Technical Analysis of the Futures Markets*, second edition, and *Intermarket Technical Analysis* are highly recommended. There are excellent books on more specific topics. Of these you should consider reading John Bollinger's *Bollinger on Bollinger Bands* and Martin Pring's *Pring on Market Momentum*. Two other more comprehensive books worth considering are Peter Bernstein's *The Portable MBA in Investment* and *The Encyclopedia of Technical Market Indicators* by Robert W. Colby and Thomas A. Meyers; the latter offers an intelligent description of the calculation and trading performance of many market indicators that could be used by traders. Comparing the results of different indicators side by side can give you valuable insight into the practical differences in these techniques. More recently, Thomas Bulkowski's *Encyclopedia of Chart Patterns* is a well-organized, clear review and analysis of all chart patterns.

The basic reference book for general contract information has always been the *Commodity Trading Manual*, but each year *Futures* magazine publishes a *Reference Guide* that gives the trading hours, contract size, and other specifications of the primary futures and options markets traded around the world. All of this information is also available on the Internet. For reviewing the basics there is Jeffrey B. Little and Lucien Rhodes, *Understanding Wall Street*, third edition, and Todd Lofton's *Getting Started in Futures*, fourth edition. The introductory material is not repeated here.

A good understanding of the most popular charting method requires reading the classic by Robert D. Edwards and John Magee (and now W. H. C. Bassetti), *Technical Analysis of Stock Trends*, eighth edition (originally published by John Magee), a comprehensive study of bar charting. For a constant flow of both classic and new techniques, the magazines *The Technical Analyst* (a U.K. publication), *Technical Analysis of Stocks & Commodities*, *Futures*, and *Active Trader* have numerous articles on trading systems and methods. A basic understanding of market phenomena and relationships, often requiring some math skill, can be found in the *Financial Analysts Journal.*

On general market lore and to provide motivation when trading is not going as well as expected, the one book that stands out is Edwin Lefevre, *Reminiscences of a Stock Operator* (originally published by Doran, reprinted by John Wiley & Sons in 1994). Richard D. Wyckoff mixes humor and philosophy in most of his books, but *Wall Street Ventures and Adventures through Forty Years* may be of general interest. Jack Schwager's *Market Wizards* is now considered a classic.

There are a number of associations and user groups that can be very helpful to traders at all levels. The Market Technician's Association (MTA), found at www.MTA.org, offers a Certified Market Technician credential, and the Association for Investment

Management Research (AIMR) offers the Charter Financial Analyst (CFA) credential. For those with higher math skills, the International Association of Financial Engineers (IAFE) offers excellent resources, and the TradeStation users groups, found in larger cities and on the Internet, can be a means for solving a difficult problem. Readers will also find Sunny Harris's *TradeStation Made Easy!* a valuable resource.

As for this book, a reader with a good background in high school mathematics can follow everything but the more complex parts. An elementary course in statistics is ideal, but knowledge of the type of probability found in Edward Thorp's *Beat the Dealer* is adequate. Fortunately, computer spreadsheet programs, such as Excel, allow anyone to use statistical techniques immediately, and most of the formulas in this book are presented in such a way that they can be easily adapted to spreadsheets. Even better, if you have a computer with trading software, such as TradeStation Technologies' TradeStation Platform, MetaStock, or any number of other products, you are well equipped to continue. If you have a live data feed, such as Bloomberg or CQG, you will also have access to technical studies that you will find very helpful. Bloomberg is also an excellent source of data.

RESEARCH GUIDELINES

Before starting, there are eight guidelines that may help make the task of developing a trading system easier.

1. **Know what you want to do before you start.** Base your trading on a sound premise. It could be an observation of how prices move in response to government policy, a theory about how prices react to economic reports, or simply a pattern that shows up at the same time each day or each month. This is the *underlying premise* of your method. It cannot be discovered by testing everything on a computer. It comes from the experience of observing price movement, reminiscent of Jesse Livermore. If that's not possible, then select ideas from credible books or articles.

2. **State your idea or premise in its simplest form.** The more complex it is, the more difficult it will be to evaluate the answer. More complex methods do not usually work as well as simple ones. Remember *Occam's razor*.

3. **Do not assume anything.** Many projects fail on basic assumptions that were incorrect. It takes practice to avoid making assumptions and to be critical of certain elements that you believe to be true. Prove everything to your own satisfaction.

4. **Try the simplest and most important parts first.** Some of the rules in your trading program will be more important than others. Try those first. It's best to understand how each rule or technique contributes to the final system. Then build slowly and carefully to prove the value of each element of the system. The ability to readily understand the operation of each part of your system is called a *transparent solution*, rather than a *fully integrated* or *complex* one. Transparent solutions are very desirable.

5. **Watch for errors of omission.** It may seem odd to look for items that are not there, but you must continually review your work, asking yourself if you have included all the necessary costs and accounted for all the risk. Simply because all the questions were answered correctly does not mean that all the right questions were asked.

6. **Question the good results.** There is a tendency to look for errors when results are extremely bad, but to accept the results that are very good. Exceptionally good results are just as likely to be caused by errors in rules, formulas, or data. They need to be checked as carefully as extremely bad results. "Surprisingly good" results are often wrong.

7. **Do not take shortcuts.** It is sometimes convenient to use the work of others to speed up the research. Check their work carefully; do not use it if it cannot be verified. Check your spreadsheet formulas manually. One error can ruin all of your hard work.

8. **Start at the end.** Define your goal and work backward to find only the necessary input. In this way, you only work with information relevant to the results; otherwise, you may expend a lot of unnecessary effort.

OBJECTIVES OF THIS BOOK

This book is intended to give you a complete understanding of the tools and techniques needed to develop or choose a trading program that has a good chance of being successful. Execution skill and market psychology are not considered—only the strategies, the methods for testing those strategies, and the means for controlling the risk. This is a goal of significant magnitude.

Not everything can be covered in a single book; therefore, some guidelines were needed to control the material included here. Every technique in this book qualifies as systematic; that is, each has clear rules. Most of them can be automated. We begin with basic concepts and definitions, such as how much data to use, how to create an index, some statistics and probability, and other tools that are used throughout the book. The next several chapters cover the techniques that are most important to trading, such as the trend and momentum. All chapters are organized by common grouping so that you can compare variations of the same basic method. Although charting is an extremely popular technique, it is included only to the degree that it can be compared with other systematic methods, or when various patterns can be used in a computerized program (such as identifying a key reversal day). There has been no attempt to provide a comprehensive text on charting; however, various formations may offer very realistic profit objectives or provide reliable entry filters.

Neither stock options nor options on futures are included in this book. Although there are strategies that combine outright trading of stocks or futures with options, the subject is too large and too specialized to be included here. There are already many good books on options strategies.

This book does not attempt to prove that one system is better than another, because it is not possible to know what will happen in the future or how each reader will cleverly apply these techniques. Instead the book evaluates the conditions under which certain methods are likely to do better and situations that will be harmful to specific approaches. By grouping similar systems and techniques together, and by presenting many of the results in a uniform way, you should be able to compare the differences and study the results. Seeing how analysts have modified existing ideas can help you decide how to proceed and give you an understanding of why you might choose one path over another. With a more complete picture, common sense should prevail over computing power.

PROFILE OF A TRADING SYSTEM

There are quite a few steps to be considered when developing a trading program. Some of these are simply choices in style, while others are essential to the success of the results. They have been listed here and are discussed briefly as items to bear in mind as you continue the process of creating or choosing a trading system.

Changing Markets and System Longevity

Markets are not static. They evolve as does everything else. During the past 10 years, changes have been predominately in technology, participation, globalization, and the cost of doing business.

Technology includes communications, trading equipment (primarily computers and handheld devices), electronic exchanges, and order entry. These innovations have accelerated the trading process, provided faster access to quotes, and instantaneous order entry for computerized strategies. Electronic markets have changed the nature of the order flow and made information about buyers and sellers more accessible. They have accelerated the price discovery process and changed the way prices react to news, and they have facilitated high frequency trading.

Increased participation that followed the historic bull market of the 1990s—aided by the proliferation of financial news networks, better communications, and faster computers—did not suffer during the economic crisis that began in 2008. More participation has changed the level of noise in individual stocks and futures, but it is most obvious in the index markets. Noise results from a large, constant flow of orders placed for unrelated reasons.

Globalization is mostly the result of the reliability of advances in communications. Not only can we see the same news at the same time everywhere in the world, but we can pass information quickly electronically. Equally important, we do not think about the reliability of our communications equipment. We expect our telephones and Internet connections, mostly wireless, to work without question. When we trade, we are willing to bet on it.

The dramatic reduction in commission cost has been a major influence on trading, opening up opportunities for the fast trader. Negotiated commissions have served the

God of Competition. For institutions, stock transactions can be done at a fraction of a cent per share, and for the general public, anyone can get $10 per order. This not only facilitates fast trading but encourages greater participation. Everyone wins.

The challenge for the trader is to find a system that will adapt to future changes, whatever they are. Most changes are not sudden, but are gradually reflected in price patterns (alternating with an occasional price shock). The steady change in the percentage of institutional volume compared to individual trader orders will slowly alter price patterns. The increase in overall participation affects the level of market noise and may also affect volatility and risk. The increase in trading choices—ETFs, mutual funds, stocks, futures, options—causes a complex interdependence of markets. Index arbitrage and the trading of sector ETFs force the component stocks to move in the same direction regardless of their individual fundamentals. The creation of your own successful trading program may be of the utmost simplicity or include strategies that adapt to an uncertain future. It is both challenging and rewarding to create a program with longevity.

The Choice of Data

System decisions are limited by the data used in the analysis. Although price and volume for the specific stock or futures market may be the definitive criteria, there is a multitude of other valid statistical information that might also be used. Some of this data is easily included, such as price data from companies in the same sector or industrial group, or the current yield curve relationship. Other statistical data, including the wide range of U.S. economic data and weekly energy inventories, may add a level of robustness to the results but are less convenient to obtain and less timely.

Diversification

Not all traders are interested in diversification, which tends to reduce returns at the same time that it limits risk. Concentrating all of your resources on a single market that you understand may produce a specialized approach and much better results than using a more generalized technique over more markets. Diversification may be gained by trading two or more unique strategies applied to the same market, instead of one strategy used on a broad set of markets.

Trade Selection

Although a trading system produces signals regularly, it is not necessary to enter all of them. Selecting one over another can be done by a method of filtering. This can be a confirmation from another technique or system, a limitation on the amount of risk that can be accepted on any one trade, the use of outside information, or the current volume. Many of these additional rules add a touch of reality to an automated process. You may find, however, that too many filters result in overfitting or no trading.

Testing

A mistake in testing may cause you to trade a losing strategy or discard a profitable one. Backtesting is the only option available to confirm or validate your ideas. Testing is misguided when it is used to "discover" a successful trading method by massive scanning of combinations of techniques. The purpose of testing is to validate an idea and show robustness—that the method works over a wide range of situations in a similar manner. It can also provide a good indication of expectations, both returns and risk. A robust solution, one that works on many stocks or across similar markets, will never appear as good as an optimized result of a single stock. But using the same system for all stocks in the same sector will give you a more realistic assessment of expectations and a much better chance of success.

Risk Control

Trading survival is based on risk control. Most analysts believe that nearly any system can be profitable with proper risk management. This also means that any system can lead to ruin without risk controls. Risk must be addressed at all levels. It begins with the individual trade, but must also reflect all trades in a common sector, the risk of the single system portfolio, and finally the risk of a portfolio of systems. Trade risk can be controlled using a stop-loss but can effectively be managed by volatility. Futures traders must also pay attention to leverage. Risk management does not need to be complex, but it has many tiers.

Transaction Costs

A system that performs well on paper may be dismal when actually traded. Part of a trading program is knowing how to enter and exit the market, as well as having realistic expectations about the transaction costs, both commissions and slippage. Short-term, fast trading systems are most sensitive to transaction costs because the expected profit on each trade is small. Directional trading strategies, those that buy as prices are rising and sell when they are falling, have larger slippage than mean reversion techniques.

There is equal damage in overstating costs as there is in underestimating them. By burdening a system with unrealistic fees, tests may show a loss instead of a profit, causing you to reject a successful trading method.

Performance Monitoring and Feedback

A system is not done when you begin trading; it is only entering into a new phase. Actual trading results must be carefully monitored and compared with expectations in order to know if the system is performing properly. It is very likely that actual execution slippage

will cause you to make some changes to the system rules or to the size of the position traded. Performance monitoring provides the essential feedback needed to be successful. Even a well-designed and well-tested program may start out badly, but proper monitoring can put it on track.

A WORD ABOUT THE NOTATION USED IN THIS BOOK

In order to make the contents of this book more useful for trading, some of the traditional mathematical formulas are also shown as a single line in Microsoft's Excel notation, as well as TradeStation's *EasyLanguage*. EasyLanguage can be understood by anyone who has experience with a programming language. This edition has greatly expanded the number of complete spreadsheets and programs. These are available on the Companion Website. This should be more convenient for the reader and allow updating where necessary. In addition, some of the more complex mathematical examples and some of the older trading systems have been removed from the text but made available on the Companion Website to readers who have a sense of history or would like to have a deeper understanding of the process

There are also more complex systems and indicators that appear in both Excel and EasyLanguage, but mostly in the latter. Although these programs have been entered and tested on TradeStation, there are occasional errors introduced during final editing. Recent market activity may also produce combinations of price movement that did not occur during testing. Readers are advised to check over the code and test it thoroughly before using it.

Computer software used to develop trading strategies may vary in the notation needed to express the simplest statistical functions. For the standard deviation, Excel uses *stdev* while EasyLanguage uses *stddev*. One program expects the mean to be *avg* while another requires *average*. Please check each formula and solution for notation consistent with your needs.

AND FINALLY . . .

Throughout this book the principle of unnecessary plurality, better known as *Occam's razor*, will be stressed. The principle states that, *given more than one explanation or solution, the simplest one is the preferred*. When developing or choosing a trading strategy, it is normally the case that adding complexity for the sake of a few extra basis points increases the potential problems and risk more than it increases returns.

Pluralitas non est ponenda sine necessitate.

William of Ockham (ca 1285–1349)

It is not the purpose of this book to test every system and draw a conclusion as to which methods are best. That conclusion is not even possible. There are countless ways to generate trading signals, and markets change over time. The goal here is to provide the tools and the understanding to help aspiring and experienced traders develop systematic ways to trade that satisfy their inherent risk preference and their investment objectives. It is unlikely that any two traders will develop the same system, but the greater their knowledge, the more likely it will be profitable.

Basic Concepts and Calculations

Economics is not an exact science: it consists merely of Laws of Probability. The most prudent investor, therefore, is one who pursues only a general course of action which is "normally" right and who avoids acts and policies which are "normally" wrong.

—L. L. B. Angas

New technology gives us a sense of security. There is data from everywhere in the world at our fingertips, programs that perform sophisticated calculations instantly, and access to anyone at any time.

As Isaac Asimov foretold, there will come a time when we will no longer know how to do the calculation for long division because miniature, voice-activated computers will be everywhere. We might not even need to be able to add; it will all be done for us. We will just assume that the answer is correct, because computers don't make mistakes.

In a small way this is happening now. Not everyone checks their spreadsheet calculations by hand to be certain they are correct before going further. Nor does everyone print the intermediate results of computer calculations to verify their accuracy. Computers don't make mistakes, but people do.

With computer software and trading platforms making price analysis easier and more sophisticated, we no longer think of the steps involved in a moving average or linear regression. A few years ago, we looked at the correlation between investments only when absolutely necessary because they were too complicated and time-consuming to calculate. It would even be difficult to know if you had made a mistake without having someone else repeat the same calculations. Now we face a different problem: If the computer does it all, we lose our understanding of why a moving average trendline differs from a linear regression. Without looking at the data, we don't see an erroneous outlier or that the stock wasn't adjusted for splits. By not reviewing each hypothetical trade, we miss seeing that the slippage can turn a profit into a loss.

To avoid losing the edge needed to create a profitable trading strategy, the basic tools of the trade are explained in this chapter. Those of you already familiar with these methods may skip over it; others need to be confident that they can perform these calculations manually even while they use a spreadsheet.

Helpful Software

In Excel, many of the functions, such as the standard deviation, are readily accessible at any time. The more advanced statistical functions require that you install the Add-Ins, which also come free with Excel. These include histograms, regression analysis, *F*-test, *t*-test, *z*-test, Fourier analysis, and various smoothing techniques. To install these add-ins in Excel 2010, go to File/Options/Add-Ins and select the *Analysis Toolpak*. You will also want the *Solver Add-in*. Once installed, which takes only a few seconds, these functions can be accessed in the Data menu at the top of the screen. You should find the Data Analysis and Solver options at the far right on the menu bar.

There are other very useful and user-friendly statistical programs, available at a wide range of sophistication and price. One of the best values is *Pro-Stat* by Poly Software (polysoftware.com). The examples in this chapter will use both Excel and Pro-Stat.

ABOUT DATA AND AVERAGING

The Law of Averages

We begin at the beginning, with the law of averages, a greatly misunderstood and mis-quoted principle. In trading, the law of averages is most often referred to when an abnormally long series of losses is expected to be offset by an equal and opposite run of profits. It is equally wrong to expect a market that is currently overvalued or overbought to next become undervalued or oversold. That is not what is meant by the law of averages. Over a large sample, the bulk of events will be scattered close to the average in such a way that the typical values overwhelm the abnormal events and cause them to be insignificant.

This principle is illustrated in Figure 2.1, where the number of average items is extremely large, and the addition of a small abnormal grouping to one side of an average group of near-normal data does not affect the balance. It is the same as being the only passenger on a jumbo jet. Your weight is insignificant to the operation of the airplane and not noticed when you move about the cabin. A long run of profits, losses, or an unusually sustained price movement is simply a rare, abnormal event that will be offset over time by the overwhelming large number of normal events. Further discussion of this and how it affects trading can be found in "Gambling Techniques—The Theory of Runs," Chapter 22.
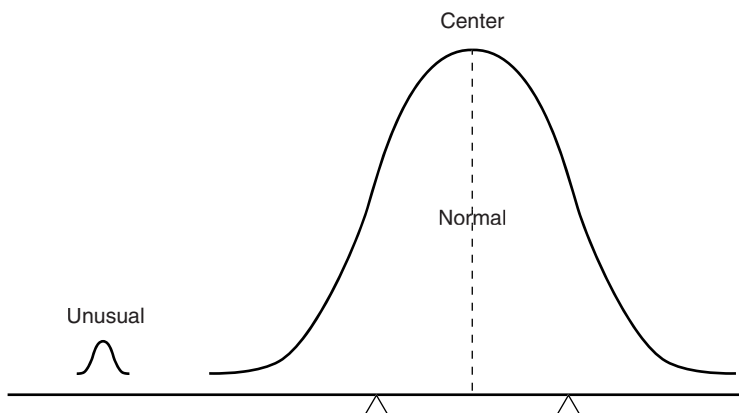
FIGURE 2.1 The Law of Averages. The normal cases overwhelm the unusual ones. It is not necessary for the extreme cases to alternate—one higher, the next lower—to create a balance.

In-Sample and Out-of-Sample Data

Proper test procedures call for separating data into *in-sample* and *out-of-sample* sets. This will be discussed in Chapter 21, System Testing. For now, consider the most important points. All testing is overfitting the data, yet there is no way to find out if an idea or system works without testing it. By setting aside data that you have not seen to use for validation, you have a better chance that your idea will work before putting money on it.

There are many ways to select in-sample data. For example, if you have 20 years of price history, you might choose to use the first 10 years for testing and reserve the second 10 years for validation. But then markets change over time; they become more volatile and may be more or less trending. It might be best to use alternating periods of in-sample and out-of-sample data, in 2-year intervals, provided that you never look at the data during the out-of-sample periods. Alternating these periods may create a problem for continuous, long-term trends, but that will be resolved in Chapter 21.

The most important factor when reserving out-of-sample data is that you get only one chance to use it. Once you have done your best to create the rules for a trading program, you then run that program through the unseen data. If the results are successful then you can trade the system, but if it fails, then you are also done. You cannot look at the reasons why it failed and change the trading method to perform better. You would have introduced *feedback*, and your out-of-sample data is considered contaminated. The second try will always be better, but it is now overfitted.

How Much Data Is Enough?

Statisticians will say, "More is better." The more data you test, the more reliable your results. Technical analysis is fortunate to be based on a perfect set of data. Each price that is

recorded by the exchange, whether it's IBM at the close of trading in New York on May 5, or the price of Eurodollar interest rates at 10:05 in Chicago, is a confirmed, precise value.

Remember that, when you use in-sample and out-of-sample data for development, you need more data. You will only get half the combinations and patterns when 50% of the data has been withheld.

Economic Data

Most other statistical data are not as timely, not as precise, and not as reliable as the price and volume of stocks, futures, ETFs, and other exchange-traded products. Economic data, such as the Producer Price Index or Housing Starts, are released as monthly averages, and can be seasonally adjusted. A monthly average represents a broad range of numbers. In the case of the PPI, some producers may have paid less than the average of the prior month and some more, but the average was +0.02. The lack of a range of values, or a standard deviation of the component values, reduces the usefulness of the information. This statistical data is often revised in the following month; sometimes those revisions can be quite large. When working with the Department of Energy (DOE) weekly data releases, you will need to know the history of the exact numbers released as well as the revisions, if you are going to design a trading method that reacts to those reports. You may find that it is much easier to find the revised data, which is not what you really need.

If you use economic data, you must be aware of when that data is released. The United States is very precise and prompt but other countries can be months or years late in releasing data. If the input to your program is monthly data and comes from the CRB Yearbook, be sure that you check when that data was actually available.

Sample Error

When an average is used, it is necessary to have enough data to make that average accurate. Because much statistical data is gathered by sampling, particular care is given to accumulating a sufficient amount of representative data. This holds true with prices as well. Averaging a few prices, or analyzing small market moves, will show more erratic results. It is difficult to draw an accurate picture from a very small sample.

When using small, incomplete, or representative sets of data, the approximate error, or accuracy, of the sample can be found using the standard deviation. A large standard deviation indicates an extremely scattered set of points, which in turn makes the average less representative of the data. This process is called the *testing of significance*. Accuracy increases as the number of items becomes larger, and the measurement of *sample error* becomes proportionately smaller

$$Sample\ error = \frac{1}{\sqrt{Number\ of\ items\ sampled}} = \frac{1}{\sqrt{N}}\ or\ \frac{1}{sqrt(N)}$$

Therefore, using only one item has a sample error of 100%; with four items, the error is 50%. The size of the error is important to the reliability of any trading system. If a system has had only four trades, whether profits or losses, it is very difficult to draw any

reliable conclusions about future performance. There must be sufficient trades to assure a comfortably small error factor. To reduce the error to 5%, there must be 400 trades. This presents a dilemma for a very slow trend-following method that may only generate two or three trades each year. To compensate for this, the identical method can be applied across many markets and the number of trades used collectively (more about this in Chapter 21).

Representative Data

The amount of data is a good estimate of its usefulness; however, the data should represent at least one bull market, one bear market, and some sideways periods. More than one of each is even better. If you were to use 10 years of daily S&P Index values from 1990 to 2000, or 25 years of 10-year Treasury notes through 2010, you would only see a bull market. A trading strategy would be profitable whenever it was a buyer, if you held the position long enough. Unless you included a variety of other price patterns, you would not be able to create a strategy that would survive a downturn in the market. Your results would be unrealistic.

Data That Is No Longer Useful

There are clear cases when a stock or futures market has undergone a structural change and the current data is different from historic data. The evolution of General Electric from a manufacturer of light bulbs to a massive financial institution represents a structural change. Its transformation back to a manufacturing company, announced in 2010, may be another structural change. A company that began in the United States, such as McDonald's, but expanded to have large international exposure, also shows a structural change in its price patterns. In foreign exchange, we have seen the individual European currencies first tied together by agreement, then finally merged into a single unit, the euro.

Is it important to include historic data in your testing when that data represents a different company profile or a different geopolitical situation? Ideally, your strategy is robust if it can adapt to these changing profiles and show consistently profitable returns over a long test period. The statisticians have that point in their favor—longer really is better. These companies and markets will continue to evolve, and your program will need to continue to adapt.

As a very fast trader, you may think about limiting your testing to much shorter periods. If you trade once each day, then in 5 years you would generate 1,250 trades; in 10 years, 2,500 trades. If your trading strategy is profitable over 2,500 trades, then you've satisfied the issue of the sampling error. However, you may not have included data that is representative of different types of markets and a variety of price patterns. Even with a large number of trades, tests spanning many years will be needed to show robustness.

Safety First

It is important to remember that the accuracy of your testing depends on both the amount of data used and the number of trades generated by the system. If your estimates of loss are not reliable, you put your investment at risk.

ON AVERAGE

In working with numbers, it is often necessary to use representative values. The range of values or the average may be substituted to change a single price into a general characteristic in order to solve a problem. The average (arithmetic mean) of many values can be a preferable substitute for any one value. For example, the average retail price of one pound of coffee in the Northeast is more meaningful to a cost-of-living calculation than the price at any one store. However, not all data can be combined or averaged and still have meaning. The average of all prices taken on the same day would not say anything about an individual market that was part of the average. Averaging the prices of unrelated items, such as a box of breakfast cereal, the hourly cost of automobile repair, and the price of the German DAX index would produce a number of questionable value. The average of a group of numbers must have some useful meaning.

The average can be misleading in other ways. Consider coffee, which rose from $0.40 to $2.00 per pound in one year. The average price of this product may appear to be $1.20; however, this would not account for the time that coffee was sold at various price levels. Table 2.1 divides the coffee price into four equal intervals, then shows that the time spent at these levels was uniformly opposite to the price rise. That is, prices remained at lower levels longer and at higher levels for shorter time periods, which is very normal price behavior.

When the time spent at each price level is included, it can be seen that the average price should be lower than $1.20. One way to calculate this, knowing the specific number of days in each interval, is by using a weighted average of the price

$$W = \frac{a_1 d_1 + a_2 d_2 + a_3 d_3 + a_4 d_4}{d_1 + d_2 + d_3 + d_4}$$

and its respective interval

$$W = \frac{6000 + 8000 + 8400 + 7200}{280}$$

$$W = 105.71$$

TABLE 2.1 Weighting an Average

Prices Go		Average During Interval	Total Days for Interval	Weighted	1/a
From	To				
40	80	$a_1 = 60$	$d_1 = 100$	6000	0.01666
80	120	$a_2 = 100$	$d_2 = 80$	8000	0.01000
120	160	$a_3 = 140$	$d_3 = 60$	8400	0.00714
160	200	$a_4 = 180$	$d_4 = 40$	7200	0.00555

This result can vary based on the number of time intervals used; however, it gives a better idea of the correct average price. There are two other averages for which time is an important element—the geometric mean and the harmonic mean.

Geometric Mean

The *geometric mean* represents a growth function in which a price change from 50 to 100 is as important as a change from 100 to 200. If there are n prices, a_1, a_2, a_3, . . . , a_n, then the geometric mean is the nth root of the product of the prices

$$G = (a_1 \times a_2 \times a_3 \times ... \times a_n)^{1/n}$$

or

```
product(a₁,a₂,a₃,...,aₙ)^(1/n)
```

To solve this mathematically, rather than using a spreadsheet, the equation above can be changed to either of two forms:

$$\ln(G) = \frac{\ln(a_1) + \ln(a_2) + ... + \ln(a_n)}{n}$$

or

$$\ln(G) = \frac{\ln(a_1 \times a_2 \times a_3 \times ... \times a_n)}{n}$$

The two solutions are equivalent. The term *ln* is the natural log, or log base *e*. (Note that there is some software where the function *log* actually is *ln*.) Using the price levels in Table 2.1,

$$\ln(G) = \frac{\ln(40) + \ln(80) + \ln(120) + \ln(160) + \ln(200)}{5}$$

Disregarding the time intervals, and substituting into the first equation:

$$\ln(G) = \frac{3.689 + 4.382 + 4.787 + 5.075 + 5.298}{5}$$

Then:

$$\ln(G) = 4.6462$$
$$G = 104.19$$

While the arithmetic mean, which is time-weighted, gave the value of 105.71, the geometric mean shows the average as 104.19.

The geometric mean has advantages in application to economics and prices. A classic example compares a tenfold rise in price from 100 to 1000 to a fall to one tenth from

100 to 10. An arithmetic mean of the two values 10 and 1000 is 505, while the geometric mean gives

$$G = (10 \times 1000)^{1/2} = 100$$

and shows the relative distribution of prices as a function of comparable growth. Due to this property, the geometric mean is the best choice when averaging ratios that can be either fractions or percentages.

Quadratic Mean

The *quadratic mean* is most often used for estimation of error. It is calculated as:

$$Q = \sqrt{\frac{\sum a^2}{N}}$$

The quadratic mean is the square root of the mean of the square of the items (*root-mean-square*). It is most well known as the basis for the standard deviation. This will be discussed later in this chapter in the section "Moments of the Distribution: Variance, Skewness, and Kurtosis."

Harmonic Mean

The *harmonic mean* is another time-weighted average, but not biased toward higher or lower values as in the geometric mean. A simple example is to consider the average speed of a car that travels 4 miles at 20 mph, then 4 miles at 30 mph. An arithmetic mean would give 25 mph, without considering that 12 minutes were spent at 20 mph and 8 minutes at 30 mph. The weighted average would give

$$W = \frac{(12 \times 20) + (8 \times 30)}{12 + 8} = 24$$

The harmonic mean is

$$\frac{1}{H} = \frac{\dfrac{1}{a_1} + \dfrac{1}{a_2} + \cdots + \dfrac{1}{a_n}}{n}$$

which can also be expressed as

$$H = n / \sum_{i=1}^{n} \left(\frac{1}{a_i} \right)$$

For two or three values, the simpler form can be used:

$$H_2 = \frac{2ab}{a+b} \qquad\qquad H_3 = \frac{3abc}{ab+ac+bc}$$

This allows the solution pattern to be seen. For the 20 and 30 mph rates of speed, the solution is

$$H_2 = \frac{2 \times 20 \times 30}{20 + 30} = 24$$

which is the same answer as the weighted average. Considering the original set of numbers again, the basic form of harmonic mean can be applied:

$$\frac{1}{H} = \frac{\frac{1}{40} + \frac{1}{80} + \frac{1}{120} + \frac{1}{160} + \frac{1}{200}}{5}$$

$$H = 87.59$$

$$= \frac{0.5708}{5} = 0.01142$$

We might apply the harmonic mean to price swings, where the first swing moved 20 points over 12 days and the second swing moved 30 points over 8 days.

PRICE DISTRIBUTION

The measurement of distribution is very important because it tells you generally what to expect. We cannot know what tomorrow's S&P trading range will be, but if the current price is 1200, then we have a high level of confidence that it will fall between 900 and 1500 this year, but less confidence that it will fall between 1100 and 1300. We have much less confidence that it will fall between 1150 and 1250, and we have virtually no chance of picking the exact range. The following measurements of distribution allow you to put a probability, or confidence level, on the chance of an event occurring.

In all of the statistics that follow, we will use a limited number of prices or—in some cases—individual trading profits and losses as the sample data. We want to measure the characteristics of our sample, finding the shape of the distribution, deciding how results of a smaller sample compare to a larger one, or how similar two samples are to each other. All of these measures will show that the smaller samples are less reliable, yet they can be still be used if you understand the size of the error or the difference in the shape of the distribution compared to the expected distribution of a larger sample.

Frequency Distributions

The *frequency distribution* (also called a *histogram*) is simple yet can give a good picture of the characteristics of the data. Theoretically, we expect commodity prices to spend more time at low price levels and only brief periods at high prices. That pattern is shown in Figure 2.2 for wheat during the past 25 years. The most frequent occurrences are at the price where the supply and demand are balanced, called *equilibrium*. When
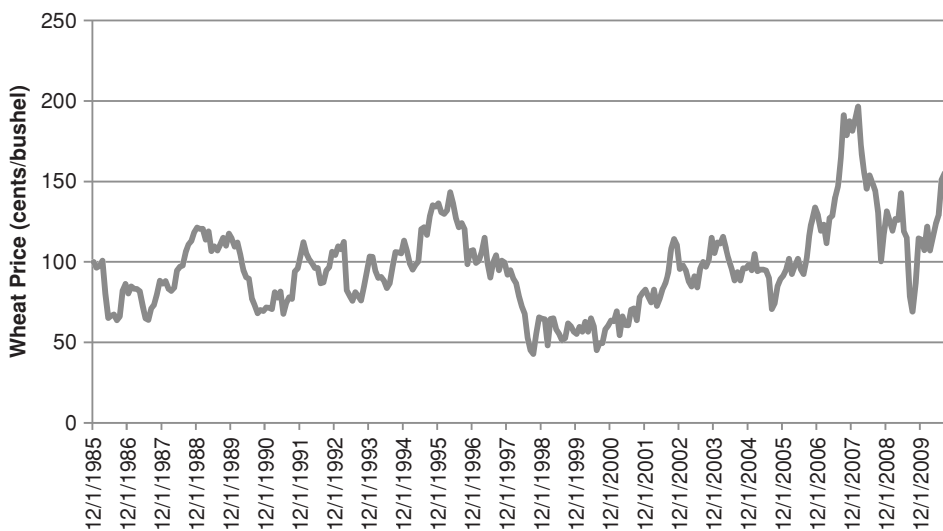
FIGURE 2.2 Wheat prices, 1985–2010.

there is a shortage of supply, or an unexpected demand, prices rise for a short time until either the demand is satisfied (which could happen if prices are too high), or supply increases to meet demand. There is usually a small tail to the left where prices occasionally trade for less than the cost of production, or at a discounted rate during periods of high supply.

To calculate a frequency distribution with 20 bins, we find the highest and lowest prices to be charted, and divide the difference by 19 to get the size of one bin. Beginning with the lowest price, add the bin size to get the second value, add the bin size to the second value to get the third value, and so on. When completed, you will have 20 bins that begin at the lowest price and end at the highest price. You then can count the number of prices that fall into each bin, a nearly impossible task, or you can use a spreadsheet to do it. In Excel, you go to Data/Data Analysis/Histogram and enter the range of bins (which you need to set up in advance) and the data to be analyzed, then select a blank place on the spreadsheet for the output results (to the right of the bins is good) and click OK. The frequency distribution will be shown instantly. You can then plot the results seen in Figure 2.3.

The frequency distribution shows that the most common price fell between $3.50 and $4.00 per bushel but the most active range was from $2.50 to $5.00. The tail to the right extends to just under $10/bushel and clearly demonstrates the *fat tail* in the price distribution. If this was a normal distribution, there would be no entries past $6. The absence of price data below $2.50 is due to the cost of production. Below that price farmers would refuse to sell at a loss; however, the U.S. government has a price support program that guarantees a minimum return for farmers.

The wheat frequency distribution can also be viewed net of inflation or changes in the U.S. dollar. This will be seen at the end of this chapter.
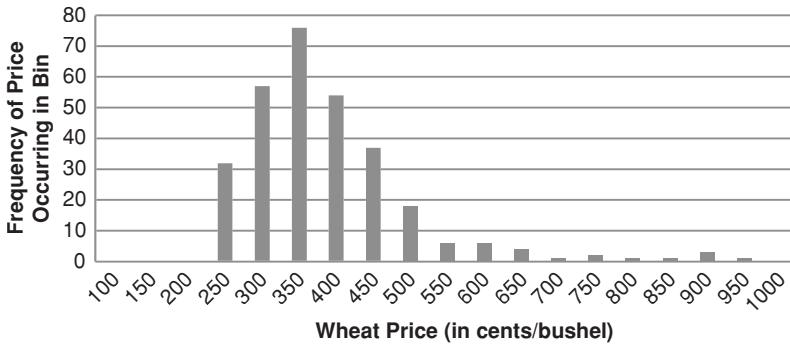
FIGURE 2.3 Wheat frequency distribution showing a tail to the right.

Short-Term Distributions

The same frequency distributions occur even when we look at shorter time intervals, although the pattern is more erratic as the time interval gets very small. If we take wheat prices for the calendar year 2007 (Figure 2.4) we see a steady move up during midyear, followed by a wide-ranging sideways pattern at higher level; however, the frequency distribution in Figure 2.5 shows a pattern similar to the long-term distribution, with the most common value at a low level and a fat tail to the right. If we had picked the few months just before prices peaked in September 2007, the chart might have shown the peak price further to the right and the fat tail on the left. For commodities, this represents a period of price instability, and expectations that prices will fall.

It should be expected that the distribution of prices for a physical commodity, such as agricultural products, metals, and energy will be *skewed* toward the left (more occurrences at lower prices) and have a *long tail* at higher prices toward the right of the chart. This is because prices remain at relatively higher levels for only short periods of time while there is an imbalance in supply and demand. In the stock market, history has shown that stocks will not sustain exceptionally high price/earnings (P/E) ratios indefinitely; however, the period of adjustment can be drawn out over many years, unlike an
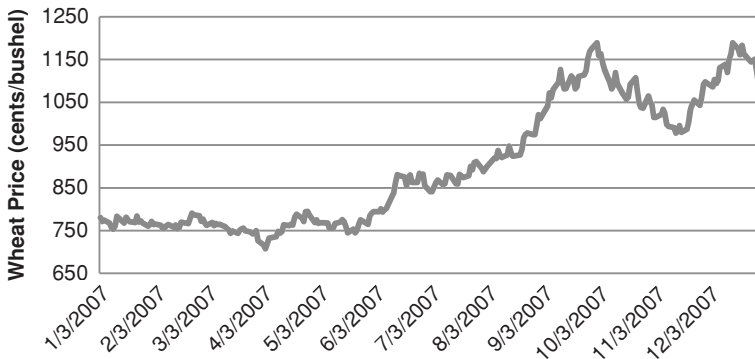
FIGURE 2.4 Wheat daily prices, 2007.

Wheat Distribution 2007
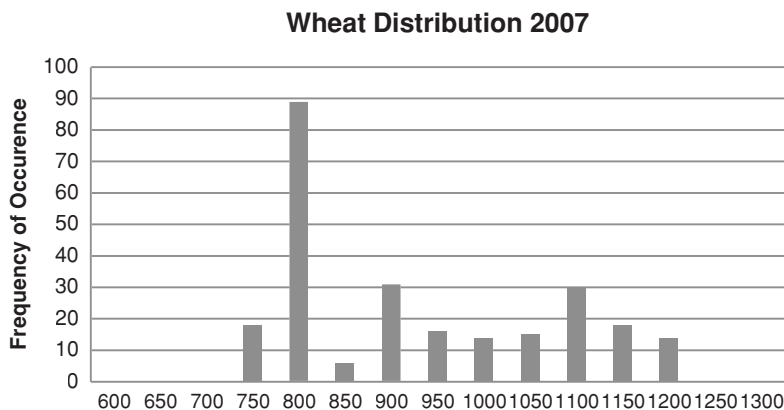
FIGURE 2.5   Frequency distribution of wheat prices, intervals of $0.50, during 2007.

agricultural product that begins again each year. When observing shorter price periods, patterns that do not fit the standard distribution may be considered in transition. Readers who would like to pursue this topic should read Chapter 18, especially the sections "Distribution of Prices" and "Steidlmayer's Market Profile."

The measures of *central tendency* discussed in the previous section are used to describe the shape and extremes of price movement shown in the frequency distribution. The general relationship between the three principal means when the distribution is not perfectly symmetric is

Arithmetic mean > Geometric mean > Harmonic mean

Median and Mode

Two other measurements, the median and the mode, are often used to define distribution. The *median*, or "middle item," is helpful for establishing the "center" of the data; when the data is sorted, it is the value in the middle. The median has the advantage of discounting extreme values, which might distort the arithmetic mean. Its disadvantage is that you must sort all of the data in order to locate the middle point. The median is preferred over the mean except when using a very small number of items.

The *mode* is the most commonly occurring value. In Figure 2.5, the mode is the highest bar in the frequency distribution, at bin 800.

In a normally distributed price series, the mode, mean, and median all occur at the same value; however, as the data becomes skewed, these values will move farther apart. The general relationship is:

Mean > Median > Mode

A normal distribution is commonly called a *bell curve*, and values fall equally on both sides of the mean. For much of the work done with price and performance data,

the distributions tend to be skewed to the right (toward higher prices or higher trading profits), and appear to flatten or cut off on the left (lower prices or trading losses). If you were to chart a distribution of trading profits and losses based on a trend system with a fixed stop-loss, you would get profits that could range from zero to very large values, while the losses would be theoretically limited to the size of the stop-loss. Skewed distributions will be important when we measure probabilities later in this chapter. There are no "normal" distributions in a trading environment.

Characteristics of the Principal Averages

Each averaging method has its unique meaning and usefulness. The following summary points out their principal characteristics:

The *arithmetic mean* is affected by each data element equally, but it has a tendency to emphasize extreme values more than other methods. It is easily calculated and is subject to algebraic manipulation.

The *geometric mean* gives less weight to extreme variations than the arithmetic mean and is most important when using data representing ratios or rates of change. It cannot be used for negative numbers but is also subject to algebraic manipulation.

The *harmonic mean* is most applicable to time changes and, along with the geometric mean, has been used in economics for price analysis. It is more difficult to calculate; therefore, it is less popular than either of the other averages, although it is also capable of algebraic manipulation.

The *mode* is the most common value and is only determined by the frequency distribution. It is the location of greatest concentration and indicates a typical value for a reasonably large sample. With an unsorted set of data, such as prices, the mode is time-consuming to locate and is not capable of algebraic manipulation.

The *median* is the middle value, and is most useful when the center of an incomplete set is needed. It is not affected by extreme variations and is simple to find; however, it requires sorting the data, which causes the calculation to be slow. Although it has some arithmetic properties, it is not readily adaptable to computational methods.

MOMENTS OF THE DISTRIBUTION: VARIANCE, SKEWNESS, AND KURTOSIS

The *moments of the distribution* describe the shape of the data points, which is the way they cluster around the mean. There are four moments: *mean*, *variance*, *skew*, and *kurtosis*, each describing a different aspect of the shape of the distribution. Simply put, the mean is the center or average value, the variance is the distance of the individual points from the mean, the skew is the way the distribution leans to the left or right relative to the mean, and the kurtosis is the peakedness of the clustering. We have already discussed the mean, so we will start with the 2nd moment.

In the following calculations, we will use the bar notation, $\bar{P}$, to indicate the average of a list of n prices. The capital P refers to all prices and the small p to individual prices.

$$\bar{P} = \frac{\sum_{i=1}^{n} p_i}{n}$$

The *mean deviation (MD)* is a basic method for measuring distribution and may be calculated about any measure of central location, such as the arithmetic mean.

$$MD = \frac{\sum_{i=1}^{n} |p_i - \bar{P}|}{n}$$

Then *MD* is the average of the differences between each price and the arithmetic mean of those prices, or some other measure of central location, with all differences treated as positive numbers. This formula will be seen often throughout the book.

Variance (2nd Moment)

Variance (Var), which is very similar to mean deviation, the best estimation of dispersion, will be used as the basis for many other calculations. It is

$$Var = \frac{\sum_{i=1}^{n} (p_i - \bar{P})^2}{n-1}$$

Notice that the variance is the square of the standard deviation, $var = s^2 = \sigma^2$, one of the most commonly used statistics. In Excel, the variance is the function *var(list)* and in TradeStation's *EasyLanguage* it is *variance(series,n)*.

The *standard deviation (s)*, most often shown as σ (sigma), is a special form of measuring average deviation from the mean, which uses the root-mean-square

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (p_i - \bar{P})^2}{n}}$$

where the differences between the individual prices and the mean are squared to emphasize the significance of extreme values, and then the total value is scaled back using the square root function. This popular measure, used throughout this book, is the *Excel* function *Stdevp* and the TradeStation function *StdDev(price,n)*, for n prices.

The standard deviation is the most popular way of measuring the dispersion of data. The value of 1 standard deviation about the mean represents a clustering of about 68% of the data, 2 standard deviations from the mean include 95.5% of all data, and 3 standard deviations encompass 99.7%, nearly all the data. While it is not possible to guarantee that all data will be included, you can use 3.5 standard deviations to include 100% of the data in a normal distribution. These values represent the groupings of a perfectly *normal* set of data, shown in Figure 2.6.
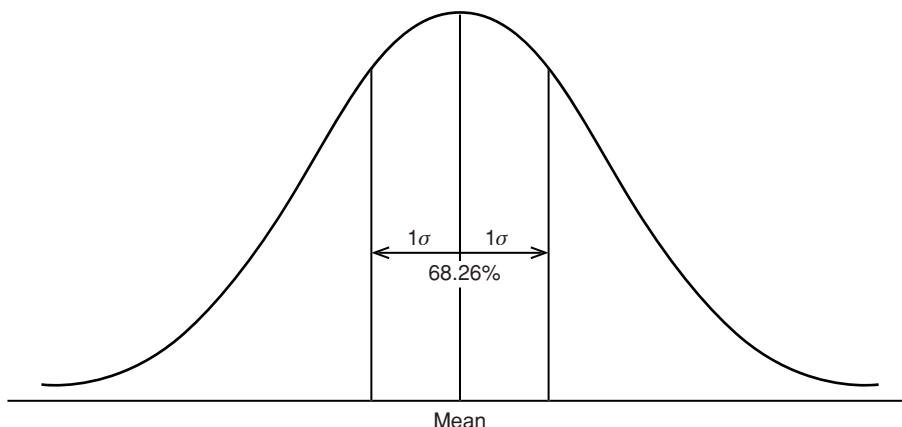
FIGURE 2.6 Normal distribution showing the percentage area included within one standard deviation about the arithmetic mean.

Skewness (3rd moment)

Most price data, however, are not normally distributed. For physical commodities, such as gold, grains, energy, and even interest rates (expressed at yields), prices tend to spend more time at low levels and much less time at extreme highs. While gold peaked at $800 per ounce for one day in January 1980, it remained between $250 and $400 per ounce for most of the next 20 years. If we had taken the average at $325 then is would be impossible for the price distribution to be symmetric. If 1 standard deviation is $140, then a normal distribution would show a high likelihood of prices dropping to $185, an unlikely scenario. This asymmetry is most obvious in agricultural markets, where a shortage of soybeans or coffee in one year will drive prices much higher, but a normal crop the following year will return those prices to previous levels.

The relationship of price versus time, where markets spend more time at lower levels, can be measured as *skewness*—the amount of distortion from a symmetric distribution, which makes the curve appear to be short on the left and extended to the right (higher prices). The extended side is called the *tail*, and a longer tail to the right is called *positive skewness*. *Negative skewness* has the tail extending toward the left. This can be seen in Figure 2.7.

In a perfectly normal distribution, the mean, median, and mode all coincide. As prices become positively skewed, typical of a period of higher prices, the mean will show the greatest change, the mode will show the least, and the median will fall in between. The difference between the mean and the mode, adjusted for dispersion using the standard deviation of the distribution, gives a good measure of skewness.

$$(\text{Skewness})\, S_K = \frac{\text{Mean} - \text{Mode}}{\text{Standard deviation}}$$
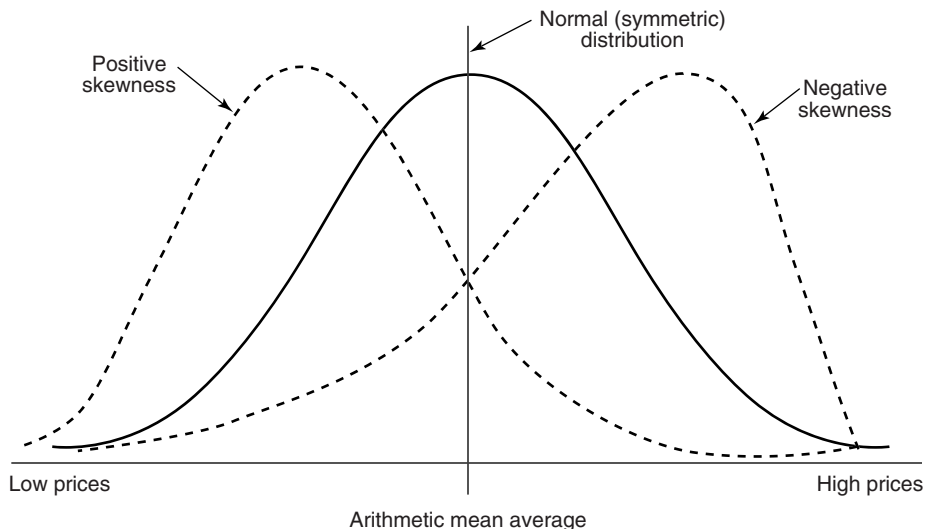
FIGURE 2.7   Skewness. Nearly all price distributions are positively skewed, showing a longer tail to the right, at higher prices.

The distance between the mean and the mode, in a moderately skewed distribution, turns out to be three times the difference between the mean and the median; the relationship can also be written as:

$$(Skewness)\,S_K = \frac{3 \times (Mean - Median)}{Standard\ deviation}$$

To show the similarity between the 2nd and 3rd moments (variance and skewness) the more common computational formula is

$$S_K = \frac{\sum_{i=1}^{n} \left(p_i - \bar{P} \right)^3}{(n-1)\sigma^3}$$

where n is the number of prices in the distribution, and σ is the standard deviation of the prices. The functions for skew can be found in Excel and TradeStation.

Transformations

The skewness of a data series can sometimes be corrected using a *transformation*. Price data may be skewed in a specific pattern. For example, if there are 3 occurrences at twice the price, and 1/9 of the occurrences at 3 times the price, the original data can be transformed into a normal distribution by taking the square root of each data item. The characteristics of price data often show a logarithmic, power, or square-root relationship.

To calculate the probability level of a distribution based on the skewed distribution of price, we can convert the normal probability to the *exponential probability* equivalent, P_E, using

$$P_E = \frac{\bar{X} \log_{10}\left(\frac{1}{1-P}\right)}{\log_{10} e}$$

where $\bar{X}$ = the average of all prices
P = the normal probability
$\log_{10} e$ = .434294482

While the normal probability, P, understates the probability of occurrence in a price distribution, the exponential distribution, P_E, will overstate the probability. Whenever possible, it is better to use the exact calculation; however, when calculating risk, it might be best to err on the side of slightly higher than expected risk.

Skewness in Distributions at Different Relative Price Levels

Because the lower price levels of most commodities are determined by production costs, price distributions show a clear tendency to resist moving below these thresholds. This contributes to the positive skewness in those markets. Considering only the short term, when prices are at unusually high levels, they can be volatile and unstable, causing a negative skewness that can be interpreted as being top heavy. Somewhere between the very high and very low price levels, we may find a frequency distribution that looks normal. Figure 2.8 shows the change in the distribution of prices over, for example, 20 days as prices move sharply higher. The mean shows the center of the distributions as they change from positive to negative skewness. This pattern indicates that a normal
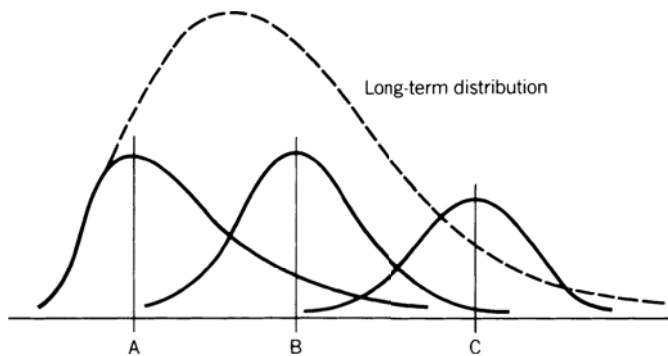
FIGURE 2.8   Changing distribution at different price levels. A, B, and C are increasing mean values of three shorter-term distributions and show the distribution changing from positive to negative skewness.

distribution is not appropriate for all price analysis, and that a log, exponential, or power distribution would only apply best to long-term analysis.

Kurtosis (4th Moment)

One last measurement, *kurtosis*, is needed to describe the shape of a price distribution. *Kurtosis* is the peakedness or flatness of a distribution as shown in Figure 2.9. This measurement is good for an unbiased assessment of whether prices are trending or moving sideways. If you see prices moving steadily higher, then the distribution will be flatter and cover a wider range. This is call *negative kurtosis*. If prices are rangebound, then the frequency will show clustering around the mean and we have *positive kurtosis*. Steidlmayer's *Market Profile*, discussed in Chapter 18, uses the concept of kurtosis, with the frequency distribution accumulated dynamically using real-time price changes.

Following the same form as the 3rd moment, skewness, kurtosis can be calculated as

$$K = \frac{\sum_{i=1}^{n} \left(p_i - \bar{P} \right)^4}{(n-1)\sigma^4}$$
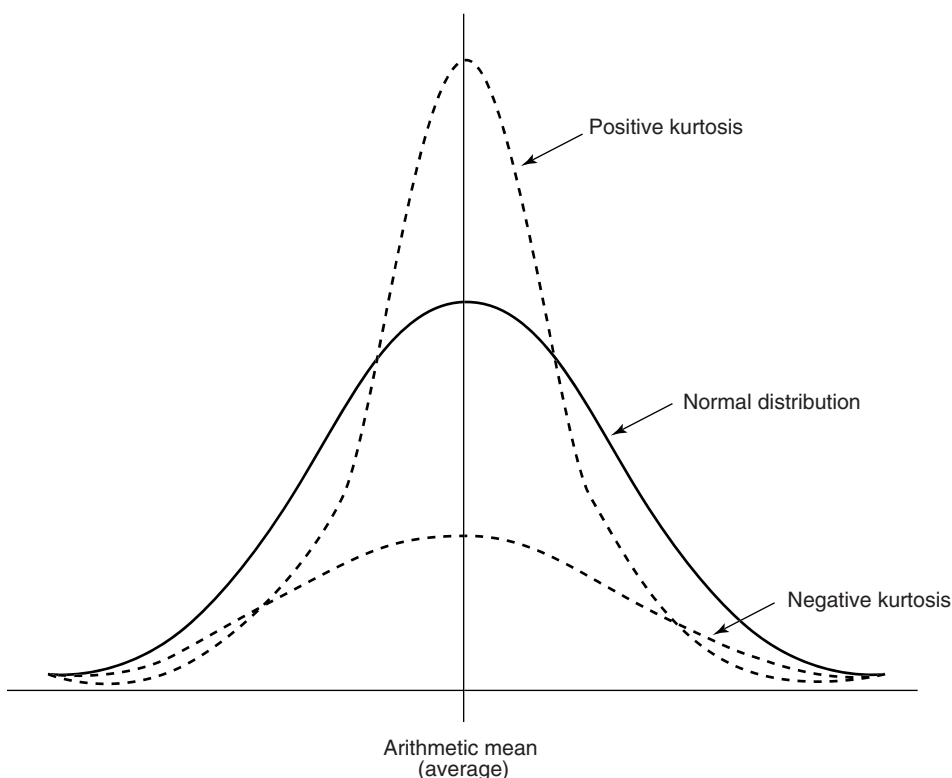
FIGURE 2.9 Kurtosis. A positive kurtosis is when the peak of the distribution is greater than normal, typical of a sideways market. A negative kurtosis, shown as a flatter distribution, occurs when the market is trending.

An alternative calculation for kurtosis is

$$K = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{p_i - \overline{P}}{\sigma} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

where n = the number of prices in the distribution
  p_i = the individual prices
  $\overline{P}$ = the average of n prices
  σ = the standard deviation of prices

Most often the *excess kurtosis* is used, which makes it easier to see abnormal distributions. Excess kurtosis, $KE = K - 3$ because the normal value of the kurtosis is 3.

Kurtosis is also useful when reviewing system tests. If you find the kurtosis of the daily returns, they should be somewhat better than normal if the system is profitable; however, if the kurtosis is above 7 or 8, then it begins to look as though the trading method is overfitted. A high kurtosis means that there are an overwhelming number of profitable trades of similar size, which is not likely to happen in real trading. Any high value of kurtosis should make you immediately suspicious.

Choosing between Frequency Distribution and Standard Deviation

Frequency distributions are important because the standard deviation doesn't work for skewed distributions, which is most common for most price data. For example, if we look back at the histogram for wheat, the average price over the past 25 years was $3.62 and the standard deviation of those prices was $1.16, then 1 standard deviation to the left of the mean is $2.46, a bin which has no data. On the right side, 3.5 standard deviations, which should contain 100% of the data, is $7.68, far below the actual high price.

Then using the standard deviation can fail on both ends of the distribution for highly skewed data, while the frequency distribution gives a very clear and useful picture. If we wanted to know the price at the 10% and 90% probability levels based on the frequency distribution, we would sort all the data from low to high. If there were 300 monthly data points, then the 10% level would be in position 30 and the 90% level in position 271. The median price would be at position 151. This is shown in Figure 2.10.

When there is a long tail to the right, both the frequency distribution and the stand-ard deviation imply that large moves are to be expected. When the distribution is very symmetric, then we are not as concerned. For those markets that have had extreme moves, neither method will tell you the size of the extreme move that could occur. There is no doubt that, given enough time, we will see profits and losses that are larger than we have seen in the past, perhaps much larger.

Autocorrelation
Serial correlation or *autocorrelation* means that there is persistence in the data; that is, future data can be predicted (to some degree) from past data. Such a quality could indicate
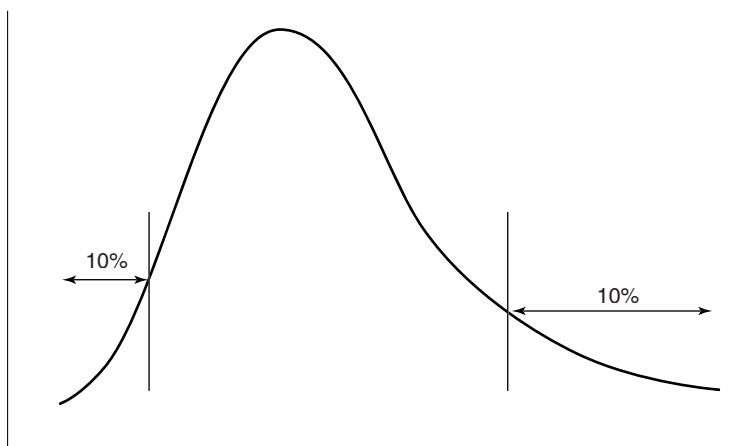
FIGURE 2.10 Measuring 10% from each end of the frequency distribution. The dense clustering at low prices will make the lower zone look narrow, while high prices with less frequent data will appear to have a wide zone.

the existence of trends. A simple way of finding autocorrelation is to put the data into column A of a spreadsheet, then copy it to column B while shifting the data down by 1 row. Then find the correlation of column A and column B. Additional correlations can be calculated shifting column B down 2, 3, or 4 rows, which might show the existence of a cycle.

A formal way of finding autocorrelation is by using the Durbin-Watson test, which gives the *d*-statistic. This approach measures the change in the errors (*e*), the difference between *N* data points and their average value.

$$e_t = r_t - \frac{\sum_{t-N+1}^{t} r_i}{N}$$

$$d = \frac{\sum_{t-N+1}^{t} \left(e_i - e_{i-1}\right)^2}{\sum_{t-N+1}^{t} e_i^2}$$

The value of *d* always falls between 0 and 4. There is no autocorrelation if *d*=2. If *d* is substantially less than 2 there is positive autocorrelation; however, if it is below 1, then there is more similarity in the errors than is reasonable. The farther *d* is above 2 the more negative autocorrelation appears in the error terms.

A *positive autocorrelation*, or serial correlation, means that a positive error factor has a good chance of following another positive error factor.

Probability of Achieving a Return

To be uncertain is to be uncomfortable, but to be certain is to be ridiculous.
—Chinese proverb

If we see the normal distribution (Figure 2.6) as the annual returns for the stock market over the past 50 years, then the mean is about 8%, and one standard deviation is 16%. In any one year, we can expect the returns to be 8%; however, there is a 32% chance that it will be either greater than 24% (the mean plus one standard deviation) or less than –8% (the mean minus one standard deviation). If you would like to know the probability of a return of 20% or greater, you must first rescale the values,

$$Probability\ of\ reaching\ objective = \frac{Objective - Mean}{Standard\ deviation}$$

If your objective is 20%, we calculate

$$Probability = \frac{20\% - 8\%}{16\%} = 0.75$$

Table A1.1, Appendix 1 gives the probability for normal curves. Looking up the standard deviation of 0.75 gives 27.34%, a grouping of 54.68% of the data. That leaves one half of the remaining data, or 22.66%, above the target of 20%.

Calculating the Probability Automatically
It is inconvenient to look up the probability values in a table when you are working with a spreadsheet or computer program, yet the probabilities are easier to understand than standard deviation values. You can calculate the area under the curve that corresponds to a particular z value (the standard deviation), using the following approximation.[1]

Let $z' = |z|$, the absolute value of z. Then

$$r = 1 + z' \times \left(c_1 + z' \times \left(c_3 + z' \times \left(c_5 + z' \times c_6\right)\right)\right)$$

Where $c_1 = 0.049867347$
$c_2 = 0.0211410061$
$c_3 = 0.032776263$
$c_4 = 0.0000380036$
$c_5 = 0.0000488906$
$c_6 = 0.000005383$

Then the probability, P, that the returns will equal or exceed the expected return is

$$P = 0.5 \times e^{[\ln(r) \times (-16)]}$$

---

[1] Stephen J. Brown and Mark P. Kritzman, *Quantitative Methods for Financial Analysis*, 2nd ed. (Dow Jones-Irwin, 1990), 238–241.

Using the example where the standard deviation $z = 0.75$, we perform the calculation

$$r = 1 + 0.75 \times (0.049867347 + 0.75 \times [0.0211410061 + 0.75 \times (0.0032776232 \\ + 0.75 \times [0.0000380036 + 0.75 \times (0.0000488906 + 0.75 \times [0.000005383])])])$$

$$r = 1.0507$$

Substituting the value of r into the equation for P, we get

$$P = 0.5 \times e^{[\ln(1.0507) \times (-16)]} = 0.226627$$

Then there is a 22.7% probability that a value will exceed 0.75 standard deviations (that is, fall on one end of the distribution outside the value of 0.75). The chance of a value falling inside the band formed by ± 0.75 standard deviations is $1 - (2 \times 0.2266) = 0.5468$, or 54.68%. That is the same value found in Table A1.1, Appendix 1.

For those using Excel, the answer can be found with the function *normdist(p,mean, stdev,cumulative)*, where

p is the current price or value

mean is the mean of the series of p's

stdev is the standard deviation of the series of p's, and

cumulative is "true" if you want the z value.

Then the result of *normdist*(35,20,5,true) is 0.99865, or a 99.8% probability, and if cumulative is "false" then the result is 0.000866.

Standard Error

Throughout the development and testing of a trading system, we want to know if the results we are seeing are as expected. The answer will always depend on the size of the data sample and the amount of variance that is typical of the data during this period.

One descriptive measure of error, called the *standard error* (*SE*), uses the variance, which gives the estimation of error based on the distribution of the data using multiple data samples. It is a test that determines how the sample means differ from the actual mean of all the data. It addresses the *uniformity* of the data.

$$SE = \sqrt{\frac{Var}{n}}$$

where Var = the variance of the sample means

n = the number of data points in the sample means

Sample means refers to the data being sampled a number of times, each with n data points, and the means of these samples are used to find the variance. In most cases, we would use a single data series and calculate the variance as shown earlier in this chapter.

t-Statistic and Degrees of Freedom

When fewer prices or trades are used in a distribution, we can expect the shape of the curve to be more variable. For example, it may be spread out so that the peak of the distribution will be lower and the tails will be higher. A way of measuring how close the sample distribution of a smaller set is to the normal distribution (of a large sample of data) is to use the *t-statistic* (also called the *student's t-test*, developed by W. S. Gossett). The *t*-test is calculated according to its degrees of freedom (*df*), which is $n - 1$, where n is the sample size, the number of prices used in the distribution.

$$t = \frac{Average\ of\ price\ changes}{Standard\ deviation\ of\ price\ changes} \times \sqrt{n}$$

The more data in the sample, the more reliable the results. We can get a broad view of the shape of the distribution by looking at a few values of t in Table 2.2, which gives the values of t corresponding to the upper tail areas of 0.10, 0.05, 0.025, 0.01, and 0.005. The table shows that as the sample size n increases, the values of t approach those of the standard normal values of the tail areas.

The values of t that are needed to be significant can be found in Appendix 1, Table A1.2, "*t*-Distribution." The column headed "0.10" gives the 90% confidence level, "0.05" is 95%, and "0.005" is 99.5%. For example, if we had 20 prices in our sample, and wanted the probability of the upper tail to be 0.025, then the value of t would need to be 2.086. For smaller samples, the value of t would be larger in order to have the same confidence.

When testing a trading system, degrees of freedom can be the number of trades produced by the strategy. When you have few trades, the results may not represent what you will get using a longer trading history. When testing a strategy, you will find a similar relationship between the number of trades and the number of parameters, or variables, used in the strategy. The more variables used, the more trades are needed to create expectations with an acceptable confidence level.

2-Sample *t*-Test

You may want to compare two periods of data to decide whether the price patterns have changed significantly. Some analysts use this to eliminate inconsistent data, but the char-

TABLE 2.2 Values of *t* Corresponding to the Upper Tail Probability of 0.025

Degrees of Freedom (*df*)	Value of *t*
1	12.706
10	2.228
20	2.086
30	2.042
120	1.980
Normal	1.960

acteristics of price and economic data change as part of the evolving process, and systematic trading should be able to adapt to these changes. This test is best applied to trading results in order to decide if a strategy is performing consistently. This is done with a 2-sample t-test:

$$t = \frac{\overline{P}_1 - \overline{P}_2}{\sqrt{\dfrac{var_1^2}{n_1} + \dfrac{var_2^2}{n_2}}}$$

where $\overline{P}_1$ and $\overline{P}_2$ = the averages of the prices for periods 1 and 2
         var_1 and var_2 = the variances of the prices for periods 1 and 2
           n_1 and n_2 = the number of prices in periods 1 and 2

and the two periods being compared are mutually exclusive. The degrees of freedom, df, needed to find the confidence levels in Table A1.2 can be calculated using Satterthwaite's approximation, where s is the standard deviation of the data values:

$$df = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{(n_1 - n_2)} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{(n_2 - n_1)}}$$

When using the t-test to find the consistency of profits and losses generated by a trading system, replace the data items by the net returns of each trade, the number of data items by the number of trades, and calculate all other values using returns rather than prices.

STANDARDIZING RISK AND RETURN

In order to compare one trading method with another, it is necessary to standardize both the tests and the measurements used for evaluation. If one system has total returns of 50% and the other 250%, we cannot decide which is best unless we know the duration of the test and the volatility of the returns, or risk. If the 50% return was over 1 year and the 250% return over 10 years, then the first one is best. Similarly, if the first return had an annualized risk of 10% and the second a risk of 50%, then both would be equivalent. The return relative to the risk is crucial to performance as will be discussed in Chapter 21, System Testing. For now it is only important that returns and risk be annualized or standardized to make comparisons valid.

Calculating Returns

The calculations for both 1-period returns and annualized returns will be an essential part of all performance evaluations. In its simplest form, the *1-period rate of return, R,* or the *holding period rate of return* is often given as

$$Return = \frac{Ending\ value - Starting\ value}{Starting\ value} = \frac{Ending\ value}{Starting\ value} - 1$$

For the stock market, which has continuous prices, this can be written

$$r_1 = \frac{p_1 - p_0}{p_0} = \frac{p_1}{p_0} - 1$$

where p_0 is the initial price and p_1 is the price after one period has elapsed. The securities industry often prefers a different calculation,

$$r_n = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

Both methods have advantages and disadvantages. Neither one is the "correct" calculation. Note that in some software, the function *log* is actually the natural log, and *log10* is the log base 10. It is best to always check the definitions. In order to distinguish the two calculations, the first method will be called the *standard method* and the second the *ln method.*

In the following spreadsheet example, shown in Table 2.3 over 22 days, the standard returns are in column D and the ln returns in column E. The differences seem small, but the averages are 0.00350 and 0.00339. The standard returns are better by 3.3% over only one trading month. At this rate, the standard method would have yielded returns that were nearly 40% higher after one year. The *net asset value* (*NAV*), used extensively throughout this book, compounds the periodic returns, and most often has a starting value, $NAV_0 = 100$.

$$NAV_t = NAV_{t-1} \times (1 + r_t)$$

Annualizing Returns

In most cases, it is best to standardize the returns by *annualizing*. This is particularly helpful when comparing two sets of test results, where each covers a different time period. When annualizing, it is important to know that

- Government instruments use a 360-day rate (based on 90-day quarters).
- A 365-day rate is common for most other data that can change daily.
- Trading returns are best with 252 days, which is the typical number of days in a trading year for the United States (262 days for Europe).

TABLE 2.3   Calculation of Returns and NAVs from Daily Profits and Losses

Date	PL	Cum PL	Standard Return	Return using LN	NAV	NAV using LN
9/10/2010		100000			100.00	100.00
9/13/2010	1154	101154	0.01154	0.01147	101.15	101.15
9/14/2010	1795	102949	0.01774	0.01759	102.95	102.91
9/15/2010	−1859	101090	−0.01806	−0.01822	101.09	101.08
9/16/2010	−1603	99487	−0.01585	−0.01598	99.49	99.49
9/17/2010	449	99936	0.00451	0.00450	99.94	99.94
9/20/2010	1090	101026	0.01090	0.01084	101.03	101.02
9/21/2010	2949	103974	0.02919	0.02877	103.97	103.90
9/22/2010	1346	105320	0.01295	0.01286	105.32	105.18
9/23/2010	64	105384	0.00061	0.00061	105.38	105.24
9/24/2010	−2051	103333	−0.01946	−0.01966	103.33	103.28
9/27/2010	3269	106602	0.03164	0.03115	106.60	106.39
9/28/2010	1795	108397	0.01684	0.01670	108.40	108.06
9/29/2010	−1154	107243	−0.01064	−0.01070	107.24	106.99
9/30/2010	128	107372	0.00120	0.00119	107.37	107.11
10/1/2010	−705	106666	−0.00657	−0.00659	106.67	106.45
10/4/2010	1090	107756	0.01022	0.01016	107.76	107.47
10/5/2010	−449	107307	−0.00416	−0.00417	107.31	107.05
10/6/2010	2308	109615	0.02150	0.02128	109.62	109.18
10/7/2010	−769	108846	−0.00702	−0.00704	108.85	108.48
10/8/2010	−256	108589	−0.00236	−0.00236	108.59	108.24
10/12/2010	−1218	107372	−0.01122	−0.01128	107.37	107.11
Average			0.00350	0.00339		
Std Dev			0.01502	0.01495		
Ann StdDev			0.23846	0.23730		
AROR					125.85%	119.69%

The following formulas use 252 days, which will be the standard throughout this book except for certain interest rate calculations; however, 365 or 360 may be substituted, or even 260 for trading days in other parts of the world.

The annualized rate of return (AROR) on a simple-interest basis for an investment over n days is

$$AROR_{\text{simple}} = \frac{E_n}{E_0} \times \frac{252}{n}$$

where E_0 is the starting equity or account balance, E_n is the equity at the end of the nth period, and $252/n$ are the years expressed as a decimal. When the 1-period returns are calculated using the standard method, then the *annualized compounded rate of return* is

$$AROR_{\text{compounded}} = \left(\frac{E_n}{E_0}\right)^{\frac{252}{n}} - 1$$

Note that *AROR* or *R* (capital) refers to the annualized rate of return while *r* is the daily or 1-period return. Also, the form of the results is different for the two calculations. An

increase of 25% for the simple return will show as 0.25 while the same increase using the compounded returns will be 1.25.

When the 1-period returns use the ln method, then the annualized rate of return is the sum of the returns divided by the number of years

$$ AROR_{\text{ln method}} = \frac{\sum_{i=1}^{n} r_i}{n} $$

An example of this can be found in column F, the row labeled AROR, in the previous spreadsheet. Note that the annualized returns using the ln method are much lower than those using division and compounding. The compounded method will be used throughout this book.

Probability of Returns

The use of the standard deviation and compounded rate of return are combined to find the probability of a return objective. In the following calculation,[2] the arithmetic mean of continuous returns is $\ln(1 + R_g)$, and it is assumed that the returns are normally distributed.

$$ z = \frac{\ln\left(\dfrac{T}{B}\right) - \ln(1 + R_g)n}{\sqrt{s \times n}} $$

where
z = standardized variable (can be found in Appendix A1)
T = target value or rate-of-return objective
B = beginning investment value
R_g = geometric average of periodic returns
n = number of periods
s = standard deviation of the logarithms of the quantities 1 plus the periodic returns

Risk and Volatility

While we would always like to think about returns, it is even more important to be able to assess risk. With that in mind, there are two extreme risks. The first is *event risk*, which takes the form of an unpredictable price shock. The worst of these is *catastrophic risk*, which will cause fatal losses or ruin. The second risk is self-induced by *overleverage*, or *gearing up* your portfolio, until a sequence of bad trades causes ruin. The risk of price shocks and leverage will both be discussed in detail later in other chapters.

The standard risk measurement is useful for comparing the performance of two systems. It is commonly applied to the returns of a single stock or an entire portfolio

---

[2] This and other very clear explanations of returns can be found in Peter L. Bernstein, *The Portable MBA in Investment.*

compared to a benchmark, such as the returns of the S&P 500 or a bond fund. The most common estimate of risk is the standard deviation, σ, of returns, r, shown earlier in this chapter. For most discussions of risk, the standard deviation will also be called *volatility*. When we refer to the *target volatility* of a portfolio, we mean the percentage of risk represented by 1 standard deviation of the returns, annualized. For example, in the previous spreadsheet, columns D and E show the daily returns. The standard deviations of those returns are shown in the same columns in the row "Std Dev" as 0.01512 and 0.01495. Looking only at column D, 1 standard deviation of 0.01502 means that there is a 68% chance of a daily profit or loss less than 1.502%. However, target volatility always refers to annualized risk, and to change a daily return to an annualized one we simply multiply by $\sqrt{252}$. Then the daily standard deviation of returns of 1.512% becomes an *annualized volatility* of 23.8%, also shown at the bottom of the spreadsheet example. Because we only care about the downside risk, there is a 16% chance that we could lose 23.8% in one year. The greater the standard deviation of returns, the greater the risk.

Beta

Beta (β) is commonly used in the securities industry to express the relationship of a single market to an index or portfolio. If beta is zero then there is no relationship; if it is positive then the single series tends to move with the index, both above and below. As beta gets larger the volatility of the single market tends to be increasingly greater than the index. Specifically,

> $0 < \beta < 1$, the volatility of the single market is less than the index
> $\beta = 1$, the volatility of the single market is the same as the index
> $\beta > 1$, the volatility of the single market is greater than the index

A negative beta is similar to a negative correlation, where the moves of the market are generally opposite to the index.

Beta is found by calculating the linear regression of the single market with the index. It is the slope of the single market divided by the slope of the index. *Alpha*, the added value, is the *y*-intercept of the solution. The values can be found using Excel, and are discussed in detail in Chapter 6. A general formula for beta is

$$Beta(A) = \text{cov}(returns\ A, returns\ B)/\text{var}(returns\ B)$$

where A is the single market and B is a portfolio or index.

Adjusting to the Target Volatility

If we have a target volatility of 12%, that is, we are willing to accept a 16% chance of losing 12% in one year, but the actual returns show an annualized volatility of 23.8% based on an investment of $100,000, then to correct to a 12% target we simply increase the investment by 23.8/12.0, or a factor of 1.98 to $198,000, while holding the same position size. This is essentially deleveraging your positions by trading a smaller percentage of your account

value. Alternatively, we can reduce the position size by dividing by 1.98 and keeping the investment the same.

All results in this book will be shown at the target volatility of 12% unless otherwise stated. This is considered a modest risk level, which can be as high as 18% for some hedge funds. It will allow you to compare various systems and test results and see them at a level of risk that is most likely to represent targeted trading results.

Annualizing Daily and Monthly Returns

The previous examples used daily data and an annualization factor of $\sqrt{252}$. For monthly data, which is most common in published performance tables, we would take the monthly returns and multiply by $\sqrt{12}$. In general, annualizing can be done by multiplying the data by the square root of the number of data items in a year. Then we use 252 for daily data, 12 for monthly, 4 for quarterly, and so on.

Monthly Data Always Appears Less Volatile

Although monthly performance results are common in financial disclosure documents, this convention works to the advantage of the person publishing the performance. It is unlikely that the highest or lowest daily *net asset value* (NAV) will occur on the last day of the month, so the extremes will rarely be seen, and the performance statistics will appear smoother than when using daily returns. Those responsible for due diligence before investing in a new product will often require daily return data in order to avoid overlooking a large drawdown that occurred mid-month.

Using the S&P index as an example, the annualized volatility of the daily returns from 1990 through 2010 was 18.6%, but based on monthly returns it was only 15.3%. The risk based on monthly returns is 17.7% lower, but the annualized rate of return would be the same because it only uses the beginning and ending values.

Downside Risk

Because the standard deviation is symmetric, a series of jumps in profits will be interpreted as larger risk. Some analysts believe that it is more accurate to measure the risk as limited to only the downside returns or drawdowns. The use of only losses is called *lower partial moments*, where *lower* refers to the downside risk, and partial means where only one side of the return distribution is used. The easiest way to see this is *semivariance*, which measures the dispersion that falls below the mean, $\overline{R}$, or some target value,

$$Semivariance = \frac{\sum_{i=1}^{n}\left(\overline{R} - r_i\right)^2}{n}, \text{ where each } r_i < \overline{R}$$

However, the most common calculation for system performance is to take the daily drawdowns, that is, the net loss on each day that the total equity is below the peak equity. For example, if the system returns had produced an equity of $25,000 on day *t*, followed by a daily loss of $500, and another loss of $250, then we would have two values as input 500/25000 and 750/25000, or 0.02 and 0.03. Only those net returns below

the most recent peaks are used in the semivariance calculation. Alternatively, you could just take the standard deviation of these daily drawdowns to find the probable size of the drawdowns.

One concern about using only the drawdowns to predict other drawdowns is that it limits the number of cases and discards the likelihood that higher-than-normal profits can be related to higher overall risk. In situations where there are limited amounts of test data, using both the gains and losses will give more robust results. When there is a large amount of data, the use of drawdowns can be a very good measurement.

A full discussion of performance measurements can be found in Chapter 21, System Testing, and also in Chapter 23 under the headings "Measuring Return and Risk" and "Ulcer Index."

THE INDEX

The purpose of an average is to transform individuality into classification. In doing that, the data is often smoothed, and useful information is gained. Indexes have attracted enormous popularity in recent years. Where there was only the Value Line and S&P 500 trading as futures markets in the early 1980s, now there are stock index futures contracts representing the markets of every industrialized country. The creation of trusts, such as SPDRs (called "Spyders," the S&P 500), Diamonds (DIA, the Dow Jones Industrials), and Qs (QQQ, the NASDAQ 100) have given stock traders a familiar vehicle to invest in the broad market rather than pick individual shares. Industrial sectors, such as pharmaceuticals, health care, and technology, first appeared as mutual funds, then as ETFs, and now can also be traded as futures. These index markets all have the additional advantage of not being constrained by having to borrow shares in order to sell short, or by the uptick rule (if it is reinstated) requiring all short sales to be initiated on an uptick in price.

Index markets allow both individual and institutional participants a number of specialized investment strategies. They can buy or sell the broad market, they can switch from one sector to another (sector rotation), or they can sell an overpriced sector while buying the broad market index (statistical arbitrage). Institutions find it very desirable, from the view of both costs and taxes, to temporarily hedge their cash stock portfolio by selling S&P 500 futures rather than liquidating stock positions. They may also hedge using options on the S&P futures or SPYs. An index simplifies the decision-making process for trading. If an index does not exist, one can be constructed to satisfy most purposes.

The index holds an important role as a benchmark for performance. Most investors believe that a trading program is only attractive if it has a better return-to-risk ratio than a portfolio of 60% stocks (as represented by the S&P 500 index), and 40% bonds (the Lehman Brothers Treasury Index). Beating the index is called *creating alpha*, proving that you're smarter than the market.

Constructing an Index

An index is standardized way of expressing price movement, normally an accumulation of percentage changes. Most indexes have a starting value of 100 on a specific date. The selection of the base year is often "convenient" but can be chosen as a period of price stability. The base year for U.S. productivity and for unemployment is 1982; for consumer confidence, 1985; and for the composite of leading indicators, 1987. The CRB Yearbook shows the Producer Price Index (PPI) from as far back as 1913. For example, the PPI, which is released monthly, had a value of 186.8 in October 2010 and 185.1 in September 2010, a 0.9184% increase in one month. An index value less than 100 means that the index has less value than when it started.

Each index value is calculated from the previous value as:

$$Current\ index\ value = Previous\ index\ value \times \left(\frac{Current\ price}{Previous\ price} \right)$$

and the 1-period returns are calculated in the same way as shown previously in this chapter.

Calculating the Net Asset Value—Indexing Returns

The last calculations shown in the spreadsheet, Table 2.3, are the *net asset value* (NAV), calculated two ways. This is essentially the returns converted to an index, showing the compounded rate of return based on daily profit and losses relative to a starting investment. In the spreadsheet, this is shown in column F using standard returns and G using ln returns.

The process of calculating NAVs can be done with the following steps:

1. Establish the initial investment, in this case $100,000, shown at the top of column C. This can be adjusted later based on the target volatility.
2. Calculate the cumulative account value by adding the daily profits or losses (column B) to the previous account value (column C).
3. Calculate the daily returns by either (a) dividing today's profit or loss by yesterday's account value to get r, or (b) taking the natural log of $1 + r$.
4. If using method (a) then each subsequent $NAV_t = NAV_{t-1} \times (1 + r)$, and if using method (b) then each $NAV_t = NAV_{t-1} + \ln(1 + r)$.

The final values of the NAV are in the last dated rows. The U.S. government requires that NAVs be calculated this way, although it doesn't specify whether returns should be based on the natural log. This process is also identical to indexing, which turns any price series into one that reflects percentage returns.

Leveraged Long or Short Index Funds

As index markets have become more popular, financial engineering has created a wide range of innovative trading vehicles. Mutual funds, such as Rydex and ProFunds, cater to market timers, a group of money managers that may trade in and out of the funds each day. These funds track the major index markets closely, but offer unique variations. There are both long and short funds, and each may be leveraged. When you buy a long fund that tracks the S&P 500 (called *Nova* by Rydex), you are simply long the equivalent of the S&P 500. However, when you buy a short S&P fund, called *Ursa*, you profit when the S&P index price drops. In addition, both Rydex and ProFunds offer leverage of 1.5 or 2.0 on these funds, so that a gain of 1.0% in the S&P 500 translates into a gain of 2.0% in ProFunds' *UltraBull S&P fund*; a drop of 1.0% in the S&P would generate a profit of 2.0% in ProFunds' *UltraBear fund*. The motivation behind the short funds, or *inverse funds*, is to circumvent the U.S. government rule that does not permit short sales in retirement accounts.

The calculation for leveraged long funds is very similar to a simple index; however, a short fund (where you profit from a decline in prices) is compounded to the upside, in the same way as a long fund. The following calculation will create a long and short index that closely approximates those used by Rydex and ProFunds. In addition, it includes the calculation of the daily high and low index values. If you intend to create a leveraged S&P index, start with the cash S&P price. Use the cash index equivalent for each of the mutual fund indexes that you plan to duplicate.

In the following calculations, *leverage* is the leverage factor of the fund. Initial index values for both long and short funds are

$$XC_1 = 100$$

$$XH_1 = XC_1 + 100 \times \left(\frac{H_1}{C_1} - 1.0 \right) \times Leverage$$

$$XL_1 = XC_1 + 100 \times \left(\frac{L_1}{C_1} - 1.0 \right) \times Leverage$$

Each subsequent index value for long funds is

$$XC_i = XC_{i-1} \times \left(\left(\frac{C_i}{C_{i-1}} - 1.0 \right) \times Leverage + 1.0 \right)$$

$$XH_i = XC_{i-1} \times \left(\left(\frac{H_i}{C_{i-1}} - 1.0 \right) \times Leverage + 1.0 \right)$$

$$XL_i = XC_{i-1} \times \left(\left(\frac{L_i}{C_{i-1}} - 1.0 \right) \times Leverage + 1.0 \right)$$

For each subsequent value for the short funds invert the middle term:

$$XC_i = XC_{i-1} \times \left(\left(\frac{C_{i-1}}{C_i} - 1.0 \right) \times Leverage + 1.0 \right)$$

$$XH_i = XC_{i-1} \times \left(\left(\frac{H_{i-1}}{C_i} - 1.0 \right) \times Leverage + 1.0 \right)$$

$$XL_i = XC_{i-1} \times \left(\left(\frac{L_{i-1}}{C_i} - 1.0 \right) \times Leverage + 1.0 \right)$$

where XC, XH, and XL = the leveraged index closing, high, and low prices
C, H, and L = the underlying close, high, and low prices or index values.

If there is no leverage, then substitute the value 1 for leverage in the equations.

Cross-Market and Weighted Index

It is very convenient to create an index for two markets that cannot normally be compared because they trade in different units. For example, if you wanted to show the spread between gold and IBM, you could index them both beginning at the same date. The new indexes would then be in the same units (percent) and would be easy to compare.

Most often, an index combines a number of related markets into a single number. A *simple aggregate index* is the ratio of unweighted sums of market prices in a specific year to the same markets in the base year. Most of the popular indexes, such as the New York Stock Exchange Composite Index, fall into this class. A *weighted aggregate index* biases certain markets by weighting them to increase or decrease their effect on the composite value. The index is then calculated as in the simple aggregate index. When combining markets into a single index value, the total of all the weights will equal 1 and all weights are expressed as a percentage.

U.S. Dollar Index

A practical example of a weighted index is the U.S. Dollar Index, traded as DX on the New York Board of Trade (NYBOT) and USDX on the Intercontinental Exchange (ICE). It is a trade-weighted geometric average of six currencies: the euro, 57.6%; the Japanese yen, 13.6%; the UK pound, 11.9%; the Canadian dollar, 9.1%; the Swedish krona, 4.2%; and the Swiss franc, 3.6%. The Dollar Index serves as a valuable economic indicator, but shows only 13.6% representing Asia. It is not a good substitute for a diversified world market portfolio.

The Dollar Index rises when the U.S. dollar increases in value relative to the other currencies. In the daily calculation of the Dollar Index, each price change is represented as a percent. If, for example, the euro rises 50 points from 1.2500 to 1.2550, the change is $1.2550/1.2500 = .004$; this is multiplied by its weighting factor 0.576 and contributes -0.002304 to the index (a rising euro is a falling dollar).

STANDARD MEASUREMENTS OF PERFORMANCE

As important as standardizing risk and return is the need to compare the performance of two funds or two trading models in order to decide which is best. That decision is normally made based on a combination of return and risk. The simplest and most practical of these measurements is the *information ratio* (IR)

$$Information\ ratio = \frac{Annualized\ returns}{Annualized\ risk}$$

where both annualized returns and annualized risk (the same as annualized volatility) have been given in the previous section of this chapter.

When you compare performance using any return/risk ratio, you are looking for the optimum point on the efficient frontier. That is, any fund with a higher return but the same risk will be preferable, and any fund with the same return but a lower risk will be preferable.

Sharpe Ratio

The *Sharpe ratio*, presented by William F. Sharpe, is the most popular of all performance measures. It differs from the more generic information ratio in that it isolates excess return by subtracting the risk-free rate of return from the fund performance

$$Sharpe\ ratio = \frac{Annualized\ returns - Risk\text{-}free\ returns}{Annualized\ risk}$$

Treynor Ratio

The *Treynor ratio* also isolates excess return; however, it replaces the annualized risk of the fund or trading program with the beta of the portfolio. Beta is the volatility (risk) of a stock relative a benchmark index, for example, the S&P. The portfolio beta is the sum of the weighted individual stock betas within the portfolio. If the fund has a beta of 1.2, then it has 20% more volatility than the overall market and moves generally in the same direction (see the previous section on *beta*). The Treynor ratio is

$$Treynor\ ratio = \frac{Annualized\ returns - Risk\text{-}free\ returns}{Program\ beta}$$

Other Performance Measurements

There are numerous variations on performance measures of varying degrees of usefulness. Most often the simplest ones are the best. The value of any measurement is to rank one trading system or fund above another in terms of risk and reward; that is, to help make the decision that one is better. The developers of each performance measure believe that each of others is flawed, yet the most popular ratios will usually rank the candidate programs in similar order.

The most common performance measure after the information ratio is simply the maximum drawdown relative the investment size. The maximum drawdown should always be measured as a percentage from a highest NAV to the lowest subsequent NAV. The maximum drawdown is important because, in a long performance record, a single, large drawdown can be lost in the standard deviation when there are an overwhelming number of "normal" drawdowns. A statistician might be satisfied saying that there is a very, very small chance of that large drawdown occurring again, but an investor might want to know that it did happen and understand why it happened. One measure that accounts for this is the *Calmar ratio*

$$Calmar\ ratio = \frac{Annualized\ return}{Maximum\ drawdown}$$

Another measure that tries to focus on the drawdowns is the *Sortino ratio*. It uses downside volatility as the risk, which is the lower partial moment of degree 2, but can also be substituted with the standard deviation of those days in which the NAV was lower than the previous high NAV (see the section on *semivariance*).

$$Sortino\ ratio = \frac{Annualized\ return - Risk\text{-}free\ return}{Downside\ volatility}$$

These performance measures will be used throughout the book when comparing different systems. They will be discussed further, along with other performance measures, in Chapter 21, System Testing.

PROBABILITY

Calculation must measure the incalculable.

—Dixon G. Watts

Change is a term that causes great anxiety. However, the effects and likelihood of a chance occurrence can only be measured—not predicted. The area of study that deals with uncertainty is probability. Everyone uses probability in daily thinking and actions. When you tell someone that you will "be there in 30 minutes," you are assuming:

- Your car will start.
- You will not have a breakdown.

- You will have no unnecessary delays.
- You will drive at a predictable speed.
- You will have the normal number of green lights.

All these circumstances are extremely probabilistic, and yet everyone makes the same assumptions. Actually, the 30-minute arrival is intended only as an estimate of the average time it should take for the trip. If the arrival time were critical, you would extend your estimate to 40 or 45 minutes to account for unexpected events. In statistics, this is called *increasing the confidence interval*. You would not raise the time to two hours because the likelihood of such a delay would be too remote. Estimates imply an allowable variation, all of which is considered normal.

Probability is the measuring of the uncertainty surrounding an average value. Probabilities are measured in percent of likelihood. For example, if M numbers from a total of N are expected to fall within a specific range, the probability P of any one number satisfying the criteria is

$$P = \frac{M}{N}, \quad 0 < P < 1$$

When making a trade, or forecasting prices, we can only talk in terms of probabilities or ranges. We expect prices to rise 30 to 40 points, or we have a 65% chance of a $400 profit from a trade. Nothing is certain, but a high probability of success is very attractive.

Laws of Probability

Two basic principles of probability are easily explained by using examples with playing cards. In a deck of 52 cards, there are 4 suits of 13 cards each. The probability of drawing a specific card on any one turn is $\frac{1}{52}$. Similarly, the chances of drawing a particular suit or card number are $\frac{1}{4}$ and $\frac{1}{13}$, respectively. *The probability of any one of these three possibilities occurring is the sum of their individual probabilities*. This is known as the *law of addition*. The probability of success in choosing a numbered card, suit, or specific card (that is, either a 10, or a spade, or the queen of hearts) is

$$p = \frac{1}{13} + \frac{1}{4} + \frac{1}{52} = \frac{18}{52} = 35\%$$

The other basic principle, the law of multiplication, states that the probability of two occurrences happening simultaneously or in succession is equal to the product of their separate probabilities. The likelihood of drawing a 3 and a club from the same deck in two consecutive turns (replacing the card after each draw) or of drawing the same cards from two decks simultaneously is

$$p = \frac{1}{13} \times \frac{1}{4} = \frac{1}{52} = 2\%$$

Joint and Marginal Probability

Price movement is not as clearly defined as a deck of cards. There is often a relationship between successive events. For example, over two consecutive days, prices must have one of the following sequences or joint events: (up, up), (down, down), (up, down), (down, up), with the joint probabilities of 0.40, 0.10, 0.35, and 0.15, respectively. In this example, there is the greatest expectation that prices will rise. The marginal probability of a price rise on the first day is shown in Table 2.4, which concludes that there is a 75% chance of higher prices on the first day and a 55% chance of higher prices on the second day.

Contingent Probability

What is the probability of an outcome "conditioned" on the result of a prior event? In the example of joint probability, this might be the chance of a price increase on the second day when prices declined on the first day. The notation for this situation (the probability of A conditioned on B) is

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{Joint\ probability\ of\ A\ and\ B}{Marginal\ probability\ of\ B}$$

then

$$P(\text{up Day 2} \mid \text{down Day 1}) = \frac{Joint\ probability\ of\ (down, up)}{Marrginal\ probability\ of\ (down\ Day\ 1)}$$

$$= \frac{0.15}{0.25} = 0.60$$

The probability of either a price increase on Day 1 or a price increase on Day 2 is

$$P(\text{either}) = P(\text{up Day 1}) + P(\text{up Day 2}) - P(\text{up Day 1 and up Day 2})$$

$$= 0.75 + 0.55 - 0.40$$
$$= 0.90$$

TABLE 2.4   Marginal Probability

		Day 2				
		Up	Down			
Day 1	Up	.40	.35	.75 Up	}	Marginal probability on Day 1
	Down	.15	.10	.25 Down		
		.55	.45			
		Up	Down			

Marginal probability on Day 2

Markov Chains

If we believe that today's price movement is based in some part on what happened yesterday, we have a situation called *conditional probability*. This can be expressed as a *Markov process*, or *Markov chain*. The results, or outcomes, of a Markov chain express the probability of a state or condition occurring. For example, the possibility of a clear, cloudy, or rainy day tomorrow can be related to today's weather.

The different combinations of dependent possibilities are given by a *transition matrix*. In our weather prediction example, a clear day has a 70% chance of being followed by another clear day, a 25% chance of a cloudy day, and only a 5% chance of rain. In Table 2.5, each possibility today is shown on the left, and its probability of changing tomorrow is indicated across the top. Each row totals 100% and accounts for all weather combinations. The relationship between these events can be shown as a continuous network (see Figure 2.11).

The Markov process can reduce intricate relationships to a simpler form. First, consider a two-state process. Using the markets as an example, what is the probability of an up or down day following an up day, or following a down day? If there is a 70% chance of a higher day following a higher day (which we can say is an uptrend) and a 55% chance of a higher day following a lower day, what is the probability of any day within an uptrend being up?

TABLE 2.5 Transition Matrix

		Tomorrow		
		Clear	**Cloudy**	**Rainy**
	Clear	0.70	0.25	0.05
Today	Cloudy	0.20	0.60	0.20
	Rainy	0.20	0.40	0.40
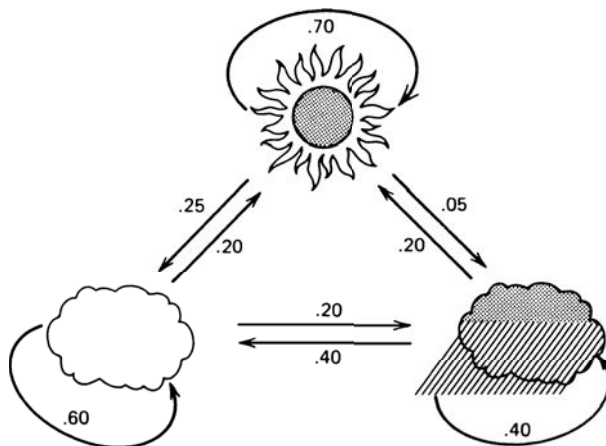
FIGURE 2.11 Probability network.

Start with either an up or down day, and then calculate the probability of the next day being up or down. This is done easily by simply counting the number of cases, given in Table 2.6a, then dividing to get the percentages, as shown in Table 2.6b.

Because the first day may be designated as up or down, it is an exception to the general rule and therefore is given the weight of 50%. The probability of the second day being up or down is the sum of the joint probabilities

$$P(\text{up})_2 = (0.50 \times 0.70) + (0.50 \times 0.55)$$

$$= 0.625$$

The probability of the second day being up is 62.5%. Continuing in the same manner, use the probability of an up day as 0.625, the down as 0.375, and calculate the third day,

$$P(\text{up})_3 = (0.625 \times 0.70) + (0.375 \times 0.55)$$

$$= 0.64375$$

and the fourth day,

$$P(\text{up})_4 = (0.64375 \times 0.70) + (0.35625 \times 0.55)$$

$$= 0.64656$$

which can now be seen to be converging. To generalize the probability of an up day, look at what happens on the ith day:

$$P(\text{up})_{i+i} = [P(\text{up})_i \times 0.70] + [(1 - P(\text{up})_i) \times 0.55]$$

Because the probability is converging, the relationship

$$P(\text{up})_{i+1} = P(\text{up})_i$$

can be substituted and used to solve the equation

$$P(\text{up})_i = [P(\text{up})_i \times 0.70] + [0.55 - P(\text{up})_i \times 0.55]$$

TABLE 2.6a Counting the Occurrences of Up and Down Days

		Today		
		Up	**Down**	**Total**
Previous day	Up	75	60	135
	Down	60	65	125

TABLE 2.6b Starting Transition Matrix

		Today		
		Up	**Down**	**Total**
Previous day	Up	0.555	0.444	1.00
	Down	0.480	0.520	1.00

giving the probability of any day being up within an uptrend as

$$P(\text{up})_i = 0.64705$$

We can find the chance of an up or down day if the 5-day trend is up simply by substituting the direction of the 5-day trend (or n-day trend) for the previous day's direction in the example just given.

Predicting the weather is a more involved case of multiple situations converging and may be very representative of the way prices react to past prices. By approaching the problem in the same manner as the two-state process, a $\frac{1}{3}$ probability is assigned to each situation for the first day; the second day's probability is

$$P(\text{clear})_2 = (0.333 \times 0.70) + (0.333 \times 0.20) + (0.333 \times 0.20)$$
$$= 0.3663$$

$$P(\text{cloudy})_2 = (0.333 \times 0.25) + (0.333 \times 0.60) + (0.333 \times 0.40)$$
$$= 0.41625$$

$$P(\text{rainy})_2 = (0.333 \times 0.05) + (0.333 \times 0.20) + (0.333 \times 0.40)$$
$$= 0.21645$$

Then, using the second day results, the third day is

$$P(\text{clear})_3 = (0.3663 \times 0.70) + (0.41625 \times 0.20) + (0.21645 \times 0.20)$$
$$= 0.38295$$

$$P(\text{cloudy})_3 = (0.3663 \times 0.25) + (0.41625 \times 0.60) + (0.21645 \times 0.40)$$
$$= 0.42791$$

$$P(\text{rainy})_3 = (0.3663 \times 0.05) + (0.41625 \times 0.20) + (0.21645 \times 0.40)$$
$$= 0.18815$$

The general form for solving these three equations is

$$P(\text{clear})_{i+i} = [P(\text{clear})_i \times 0.70] + [P(\text{cloudy})_i \times 0.20] + [P(\text{rainy})_i \times 0.20]$$

$$P(\text{cloudy})_{i+i} = [P(\text{clear})_i \times 0.25] + [P(\text{cloudy})_i \times 0.60] + [P(\text{rainy})_i \times 0.40]$$

$$P(\text{rainy})_{i+i} = [P(\text{clear})_i \times 0.05] + [P(\text{cloudy})_i \times 0.20] + [P(\text{rainy})_i \times 0.40]$$

where each $i + 1$ element can be set equal to the corresponding ith value. There are then three equations in three unknowns, which can be solved directly or by matrix multiplication, as shown in Appendix 2, "Solution to Weather Probabilities Expressed as a Markov Chain."[3] Otherwise, it will be necessary to use the additional relationship

$$P(\text{clear})_i + P(\text{cloudy})_i + P(\text{rainy})_i = 1.00$$

---

[3] A full mathematical treatment of Markov chains can be found in John G. Kemeny and J. Laurie Snell, *Finite Markov Chains* (New York: Springer-Verlag, 1976).

The results are

$$P(\text{clear}) = 0.400$$

$$P(\text{cloudy}) = 0.425$$

$$P(\text{rainy}) = 0.175$$

Bayes' Theorem

Although historic generalization exists concerning the outcome of an event, a specific current market situation may alter the probabilities. *Bayes' theorem* combines the *original probability* estimates with the *added-event probability* (the reliability of the new information) to get a *posterior* or *revised probability*:

$$\frac{P(Original\ and\ added\text{-}event)}{P(Added\text{-}event)}$$

Assume that the price changes $P(\text{up})$ and $P(\text{down})$ are both original probabilities, and an added-event probability, such as an unemployment report, trade balance, crop report, inventory stocks, or Federal Reserve interest rate announcement is expected to have an overriding effect on tomorrow's movement. Then the new probability $P(\text{Up} \mid \text{added-event})$ is:

$$\frac{P(Up\ and\ added\text{-}event)}{P(Up\ and\ added\text{-}event) + P(Down\ and\ added\text{-}event)}$$

where *up* and *down* refer to the original historic probabilities, and $P(A \text{ and } B)$ is a joint probability.

Bayes' theorem finds the conditional probability even if the joint and marginal probabilities are not known. The new probability $P(\text{up} \mid \text{added-event})$ is:

$$\frac{P(Up) \times P(Added\text{-}event \mid up)}{P(Up) \times P(Added\text{-}event \mid up) + P(Down) \times P(Added\text{-}event \mid down)}$$

where $P(\text{Added-event} \mid \text{up}) = $ the probability of the new event being a correct predictor of an upwards move

$P(\text{Added-event} \mid \text{down}) = $ the probability of prices going down when the added news indicates up

For example, if a quarter percent decline in interest rates has a 90% chance of causing stock prices to move higher, then

$$P(\text{Added-event} \mid \text{up}) = 0.90$$

and

$$P(\text{Added-event} \mid \text{down}) = 0.10$$

SUPPLY AND DEMAND

Price is the balancing point of supply and demand. In order to estimate the future price of any product or explain its historic patterns, it will be necessary to relate the factors of supply and demand and then adjust for inflation, technological improvement, and other indicators common to econometric analysis. The following sections briefly describe these factors.

Demand

The demand for a product declines as price increases. The rate of decline is always dependent on the need for the product and its available substitutes at different price levels. In Figure 2.12a, D represents normal demand for a product over some fixed period. As prices rise, demand declines fairly rapidly. D' represents increased demand, resulting in higher prices at all levels.

Figure 2.12b represents the actual demand relationship for potatoes from 1929 to 1939. Although this example is the same as the theoretical relationship in Figure 2.12a, in most cases the demand relationship is not a straight line. Production costs and minimum demand prevent the curve from going to zero; instead, it approaches a minimum price level. This can be seen previously in the frequency distribution for wheat, Figure 2.3, where the left side of the distribution falls (lower price) off sharply. On the higher end of the scale, there is a lag in the response to increased prices and a consumer reluctance to reduce purchasing even at higher prices (called "inelastic demand"). Coffee is well-known for having inelastic demand—most coffee drinkers will pay the market price rather than consume less. Figure 2.12c shows a more representative demand curve, including extremes, where 100 represents the cost of production for a producer. The demand curve,
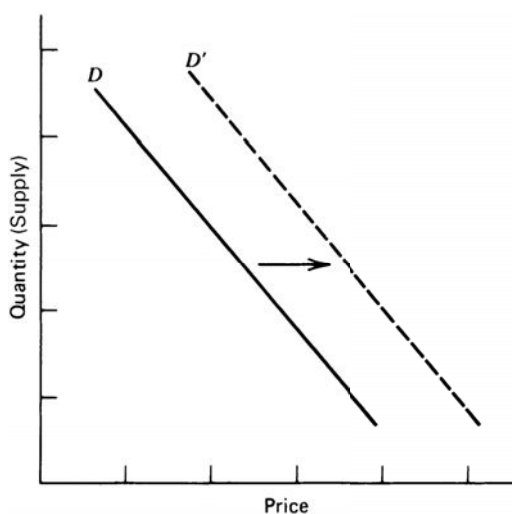
FIGURE 2.12a   Shift in demand.
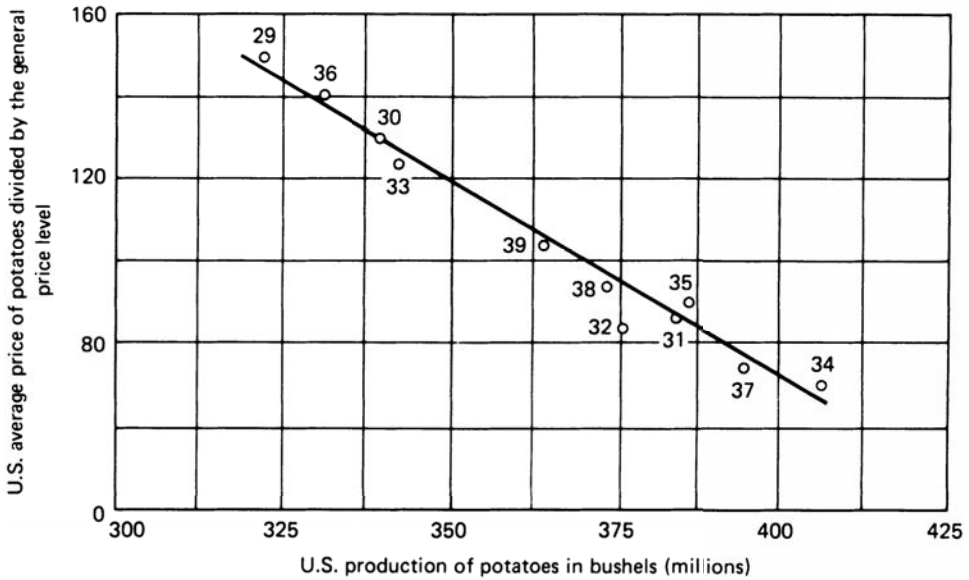
FIGURE 2.12b   Potatoes: U.S. average farm price on December 15th versus total production: 1929–1939.

therefore, shows the rate at which a change in quantity demanded brings about a change in price. Note that, although a producer may lose money below 100, lack of demand and the need for income can force sales at a loss.
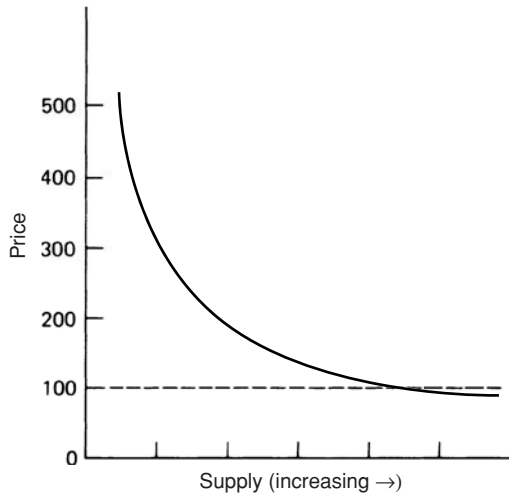
FIGURE 2.12c   Demand curve, including extremes.

Source: (for Figures 2.12a, b, and c): Geoffrey S. Shepherd and G. A. Futrell, *Agricultural Price Analysis* (Ames: Iowa State University, 1969), 53.

Elasticity of Demand

Elasticity is the key factor in expressing the relationship between price and demand and defines the shape of the curve. It is the relative change in demand as price increases:

$$E_D = \frac{Relative\ change(\%)in\ demand}{Relative\ change(\%)in\ price}$$

A market that always consumes the same amount of a product, regardless of price, is called *inelastic*; as price rises, the demand remains the same, and E_D is negatively very small. An elastic market is just the opposite. As demand increases, price remains the same and E_D is negatively very large. Figure 2.13 shows the demand curve for various levels of demand elasticity.

If supply increases for a product that has existed in short supply for many years, consumer purchasing habits will require time to adjust. The demand elasticity will gradually shift from relatively inelastic (Figure 2.13b) to relatively elastic (Figure 2.13a).

Supply

The supply side of the economic equation is the normal counterpart of demand. Figure 2.14a shows that, as price increases, the supplier will respond by offering greater amounts of the product. Figure 2.14b demonstrates the supply at price extremes. At low levels, below production costs, there is a nominal supply by those producers who must maintain operations due to high fixed costs and difficulty restarting after a shutdown (as in mining). At high price levels, supply is erratic. There may be insufficient supply in the short term, followed by the appearance of new supplies or substitutes, as in the case of a location shortage. When there is a shortage of orange juice, South American countries are willing to fill the demand; when there is an oil disruption, other OPEC nations will increase production. In most cases, however, it is reduced demand that brings price down.

Elasticity of Supply

The *elasticity of supply* E_S is the relationship between the change in supply and the change in price:

$$E_S = \frac{Relative\ change(\%)in\ supply}{Relative\ change(\%)in\ price}$$
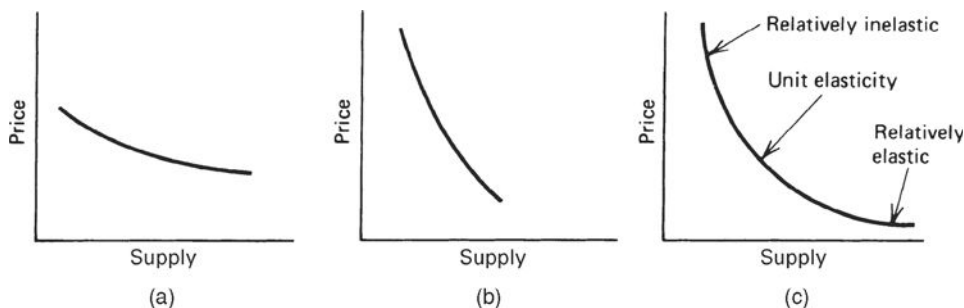
FIGURE 2.13   Demand elasticity: (a) Relatively elastic; (b) Relatively inelastic; (c) Normal market.
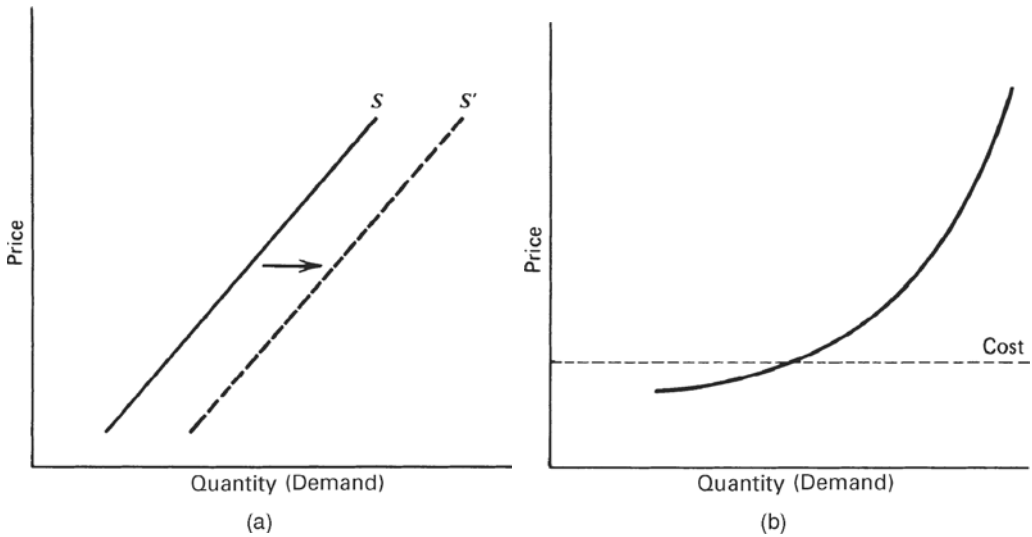
FIGURE 2.14 Supply-price relationship: (a) Shift in supply; (b) Supply curve, including extremes.

The elasticity of supply, the counterpart of *demand elasticity*, is a positive number because price and quantity move in the same direction at the same time.

Equilibrium

The demand for a product and the supply of that product cross at a point of *equilibrium*. The current price of any product, or any security, represents the point of equilibrium for that product at that moment in time. This is the basis for the technical assessment that the price, at any moment in time, represents the netting of all fundamental information. Figure 2.15 shows a constant demand line *D* and a shifting supply, increasing to the right from *S* to *S'*.

The demand line *D* and the original supply line *S* meet at the equilibrium price *P*; after the increase in supply, the supply line shifts to *S'*. The point of equilibrium *P'* represents a lower price, the consequence of larger supply with unchanged demand. Because supply and demand each have varying elasticities and are best represented by curves, the point of equilibrium can shift in any direction in a market with changing factors.

Equilibrium will be an important concept in developing trading strategies. Although the supply and demand balance may not be calculated, in practical terms equilibrium is a balance between buyers and sellers, a price level at which everyone is willing to trade, although not always happy to do so at that price. Equilibrium is associated with lower volatility and often lower volume because the urgency to buy or sell has been removed. Imbalance in the supply-demand-price relationship causes volatility. Readers interested in a practical representation of equilibrium, or price-value relationships, should study "Price Distribution Systems" in Steidlmayer's *Market Profile*, Chapter 18.
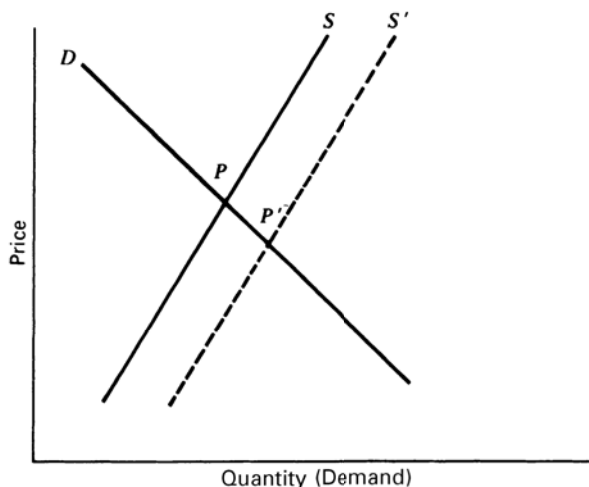
FIGURE 2.15   Equilibrium with shifting supply.

Cobweb Charts

The point at which the supply and demand lines cross is easily translated into a place on
a price chart where the direction is sideways. The amount of price volatility during this
sideways period (called noise) depends upon the price level, market participation, and
various undertones of instability caused by other factors. Very little is discussed about
how price patterns reflect the shift in sentiment between the supply and demand lines,
yet there is a clear representation of this action using cobweb charts.

Figure 2.16a shows a static (symmetric) supply-demand chart with dotted lines rep-
resenting the "cobweb." A shift in the perceived importance of supply and demand fac-
tors can cause prices to reflect the pattern shown by the direction of the arrows on the
cobweb, producing the sideways market represented by Figure 2.16b. If the cobweb were
closer to the intersection of the supply and demand lines, the volatility of the sideways
price pattern would be lower; if the cobweb were further away from the intersection, the
pattern would be more volatile.[4]

Most supply/demand relationships are not static and can be represented by lines that
cross at oblique angles. In Figure 2.17a, the cobweb is shown to begin near the intersec-
tion and move outwards, each shift forming a different length strand of the web, moving
away from equilibrium. Figure 2.17b shows that the corresponding price pattern is one
that shifts from equilibrium to increasing volatility. A reversal in the arrows on the cob-
web would show decreasing volatility moving toward equilibrium.

---

[4] Curtis McKallip, Jr., "Fundamentals behind Technical Analysis," *Technical Analysis of Stocks &
Commodities* (November 1989).

Static Supply/Demand

(a)

(b)
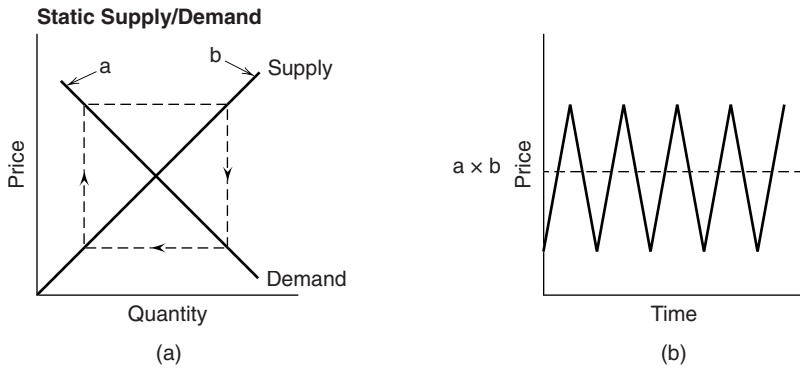
FIGURE 2.16 Static supply/demand cobweb: (a) Dotted lines represent a shift of sentiment from supply to demand to supply, and so forth; (b) the price pattern likely to result from the cobweb in (a).

Source: Curtis McKallip, Jr., "Fundamentals Behind Technical Analysis," *Technical Analysis of Stocks & Commodities* 7, no. 11 (November 1989). © 1989 by Technical Analysis, Inc. Used with permission.

Building a Model

A model can be created to *explain* or *forecast* price changes. Most models explain rather than forecast. Explanatory models analyze sets of data at concurrent times: that is, they look for relationships between multiple factors and their effect on price at the same moment in time. They can also look for *causal*, or lagged relationships, where prices *respond* to other factors after one or more days. It is possible to use the explanatory model to determine the normal price at a particular moment. Although not considered

Dynamic Supply/Demand

(a)

(b)
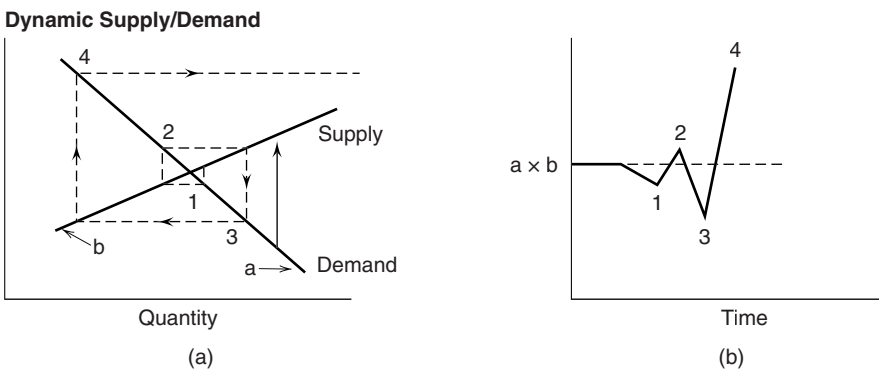
FIGURE 2.17 Dynamic supply/demand cobweb: (a) Dotted lines represent a cobweb moving away from equilibrium: (b) the price pattern shows increasing volatility.

Source: Curtis McKallip, Jr., "Fundamentals Behind Technical Analysis," *Technical Analysis of Stocks & Commodities* 7, no. 11 (November 1989). © 1989 by Technical Analysis, Inc. Used with permission.

forecasting, any variation in the actual market price from the normal or expected price could present a trading opportunity.

Methods of selecting the best forecasting model can affect its credibility. An *analytic* approach selects the factors and specifies the relationships in advance. Tests are then performed on the data to verify the premise. Many models, though, are refined by *fitting* the data, using regression analysis or some mass testing process, which applies a broad selection of variables and weighting factors to find the best fit. These models, created with perfect hindsight, are far less likely to be successful at forecasting future price levels. Even an analytic approach that is subsequently fine-tuned could be in danger of losing its forecasting ability.

The factors that comprise a model can be both numerous and difficult to obtain. Figure 2.18 shows the interrelationship between factors in the cocoa industry. Although this chart is comprehensive in its intramarket relationships, it does not emphasize the global influences that have become a major part of price movement since the mid-1970s. The change in value of the U.S. dollar and the volatility of interest rates have had far greater influence on price than some of the "normal" fundamental factors for many commodities. Companies with high debt may find the price fluctuations in their stock are larger due to interest rate changes than increases or decreases in revenues.

Models that explain price movements must be constructed from the primary factors of supply and demand. A simple example for estimating the price of fall potatoes[5] is

$$P/PPI = a + bS + cD$$

where
P = the average price of fall potatoes received by farmers
PPI = the Producer Price Index
S = the apparent domestic free supply (production less exports and diversions)
D = the estimated deliverable supply
a, b, and c = constants determined by regression analysis

This model implies that consumption must be constant (i.e., inelastic demand); demand factors are only implicitly included in the estimated deliverable supply. Exports and diversion represent a small part of the total production. The use of the *PPI* gives the results in *relative* terms based on whether the index was used as an *inflator* or *deflator* of price.

A general model, presented by Weymar,[6] may be written as three behavior-based equations and one identity:

Consumption

$$C_t = f_C(P_t, P_t^L) + e_{C_t}$$

[5] J. D. Schwager, "A Trader's Guide to Analyzing the Potato Futures Market," *1981 Commodity Yearbook* (New York: Commodity Research Bureau, 1981).

[6] F. H. Weymar, *The Dynamics of the World Cocoa Market* (Cambridge: MIT Press, 1968).

Key
Variable *A* is exogenous.
Variable *B* is determined by variables *A* and *C*.
Variable *C* is determined by variable *B*.
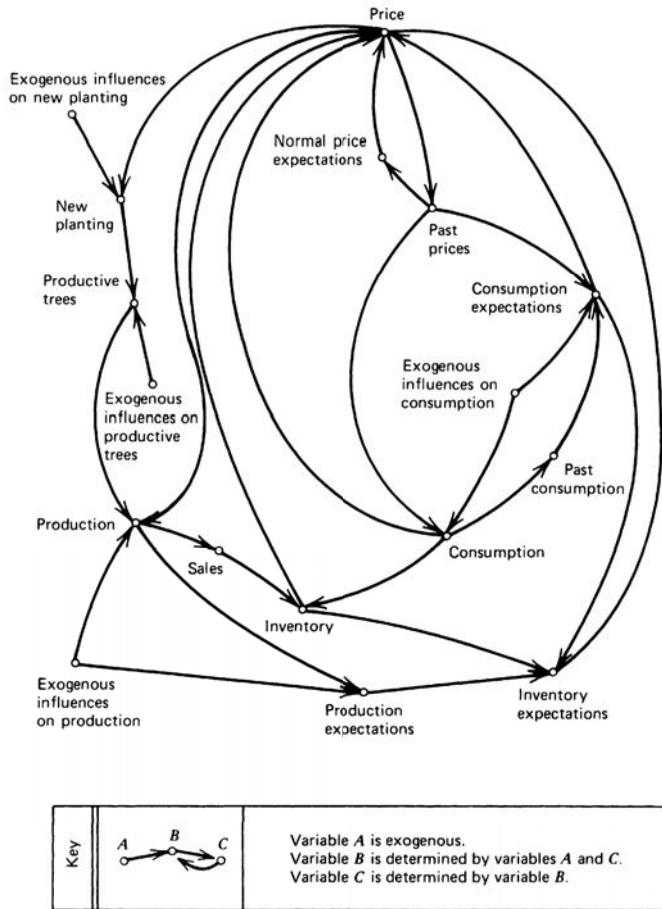
FIGURE 2.18 Cocoa factors.

Source: F. H. Weymar, *The Dynamics of the World Cocoa Market* (Cambridge: MIT Press, 1968), 2. Use by permission of The MIT Press.

Production

$$H_t = f_H(P_t, P_t^L) + e_{H_t}$$

Inventory

$$I_t = I_{t-1} + H_t - C_t$$

Supply of storage

$$P_t' - P_t = f_p(I_t) + e_p$$

where C = the consumption
P = the price

P^L = the lagged price

H = the production (harvest)

I = the inventory

P' = the expected price at some point in the future

e = the corresponding error factor

The first two equations show that both demand and supply depend on current and/ or lagged prices, the traditional macroeconomic theory; production and consumption are therefore dependent on past prices. The third equation, *inventory level*, is simply the total of previous inventories, plus new production, less current consumption. The last equation, *supply of storage*, demonstrates that people are willing to carry larger inventories if they expect prices to increase substantially. The inventory function itself, the third equation, is composed of two separate relationships: manufacturers' inventories and speculators' inventories. Each reacts differently to expected price change.

Changing Factors

Although the PPI was always considered the component of inflation and used in forecasting prices, the value of the U.S. dollar has not been an input. Currency values have always fluctuated but have taken on more significance following the dropping of the gold standard by most countries. The value of a currency is based on the health of the economy as measured by production output and inflation, among other factors.

Wheat is a good example to show the impact of these changes. As an export market for the United States, the price of wheat reflects the world value, that is, what other countries are willing to pay. Wheat is *fungible*, in other words, a country in need will buy from any source with the lowest price, and that keeps all prices competitive.

Rather than the complex analysis of factors shown in the previous section on cocoa, we will only look at the impact of inflation and currency changes on the price of wheat. Figure 2.19 shows the monthly price of cash wheat, along with the PPI and the U.S. dollar index (DX). The dollar index shows the relative value of the U.S. dollar; therefore, a decline in DX indicates a weaker U.S. dollar. Both the PPI and DX have been indexed to begin 1985 with the value of 100.

In Figure 2.19 the PPI nearly doubles from 1.0 to 1.8 while the U.S. dollar halves to 0.51. At the same time wheat prices rise 69% from 100 to 169 (the cash price from $3.56 to $7.13). If we are only concerned with the big picture, the macro factors rather than the seaonality of price, then the rise in wheat prices over this 25-year interval can be completely explained by either the PPI or the DX. In fact, they are clearly related because inflation (PPI) has a negative effect on the currency value.

By dividing the indexed price of wheat by the PPI or multiplying wheat by the DX value, Figure 2.20 shows that the long-term price of wheat is unchanged. By observation, we can conclude that the rise in wheat prices is attributed to inflation, the decline in the value of the U.S. dollar, or both.

Wheat with Inflation Components
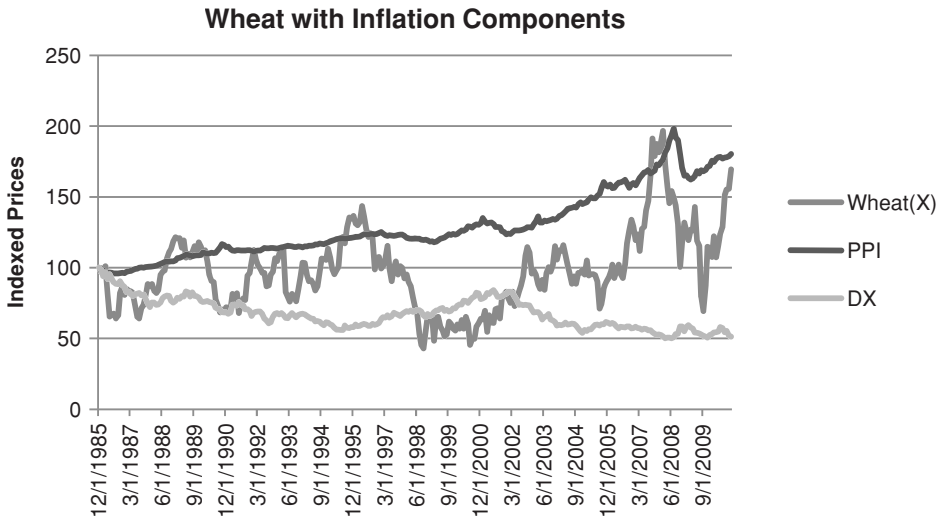
FIGURE 2.19 Cash wheat with the PPI and dollar index (DX), from 1985 through October 2010.

If we then want to look at the effect of seasonality, or supply and demand, factors that cause wheat prices to lose 75% of its value or gain 100%, we need to remove the effects of inflation first. For trading, this means selling DX in the same dollar value as buying wheat so that, if the price of wheat rises due to the U.S. dollar falling, the change in value of the hedge position is close to zero.

Effect of Inflation on Wheat Prices
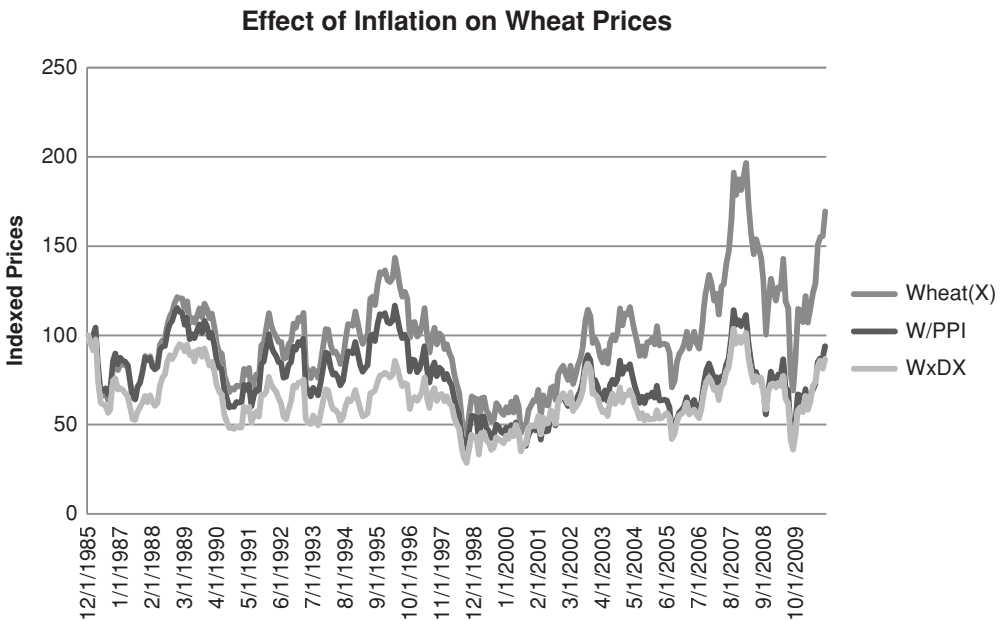
FIGURE 2.20 Wheat prices adjusted for PPI and Dollar Index.

In the case of wheat, it is clear that the price is affected by inflation. It may be that the success of most wheat trading strategies depends on the trend of inflation, rather than the underlying wheat factors. However, if the intent is to profit from seasonal moves, then a rally in May when a good crop is being harvested may be obscured by a rise in the U.S. dollar. Without hedging the potential effects of the dollar you leave the results of the strategy to chance. This will be discussed with examples in Chapter 10, Seasonality and Calendar Patterns.

Economic Reports

Economic reports are released nearly every day. Based on the economic condition of the nation, investor focus shifts from one report to another. Since 2008 the focus has been on the employment reports, the ultimate solution to recovery; GDP, a measure of that recovery; and of less importance housing, consumer confidence, and various manufacturing data. The Leading Economic Index (LEI), released by the Conference Board each month, tries to anticipate the direction of the economy about six months ahead based on

- Average weekly hours, manufacturing
- Average weekly initial claims for unemployment insurance
- Manufacturers' new orders, consumer goods and materials
- Index of supplier deliveries—vendor performance
- Manufacturers' new orders, nondefense capital goods
- Building permits, new private housing units
- Stock prices, 500 common stocks
- Money supply, M2
- Interest rate spread, 10-year Treasury bonds less federal funds
- Index of consumer expectations

Most of these seem reasonable, but the weighting of them is not clear. It has been said that the direction of the stock market plays a relatively large part in the index.

Can these and other indicators be used for trading? Are they timely, or are the expectations of their impact in the market even before the reports are released? With many reports, the market anticipates the numbers. If unemployment was expected to increase, then the stock market tends to sell off ahead of the report, or if economists anticipate the Fed lowering rates, then the yield curve will adjust to that expectation ahead of the announcement. Therefore, it is the difference between the expectation and the actual report that moves the market, and only secondarily is it the actual numbers released. For example, if the GDP was expected to rise from 3.5% to 4.0% and the actual number came in at 3.6% the market would sell off. But then it would rally again because 3.6% is still a good number indicating growth. While large, unexpected changes move the market, the cumulative effect of small changes could also be significant.

Ruggiero has quantified the significance of some of these indicators, concentrating on predicting the direction of yields, which is key to much of the financial market price moves.[7]

Interest rates

$$Ratio = 1 - \frac{Rate \ of \ Inflation}{Yield}$$

$$Inflation \ yield \ Oscilator \ (IYO) = R - \bar{R}$$

where *Yield* is the 3-month Treasury bill and $\bar{R}$ is the 20-day average of the ratio, *R*.

- If ($R < 0.2$ or IYO < 0) and $Yield_t > Yield_{t-3mo}$ then rates will rise
- If ($R > 0.3$ or IYO > 0.5) and $Yield_t < Yield_{t-3mo}$ then rates will fall

Money Supply

Using monthly data for *M2* and the 3-month Treasury bill yields, where *m* is the current month,

If ($M2_m - M2_{m-1}$) > ($M2_m - M2_{m-6}$) and $Yield_t > Yield_{t-11mo}$ then rates will rise
If ($M2_m - M2_{m-1}$) < ($M2_m - M2_{m-6}$) and $Yield_t < Yield_{t-11mo}$ then rates will fall

Consumer Sentiment

Using the University of Michigan's *Consumer Sentiment Survey* (*CS*), and where *m* is the month it is released,

If $CS_m > CS_{m-12}$ and $CS_m > CS_{m-11}$ and $Yield_t > Yield_{t-4mo}$ then rates will rise
If $CS_m < CS_{m-12}$ and $CS_m < CS_{m-11}$ and $Yield_t < Yield_{t-4mo}$ then rates will fall

Unemployment Claims

Using monthly *unemployment claims* (*UC*) released on the first Friday of each month,

If $UC_m < UC_{m-11}$ and $UC_m > UC_{m-14}$ then rates will rise
If $UC_m > UC_{m-11}$ and $UC_m < UC_{m-14}$ then rates will fall

The big picture of price direction is very important, and an accurate forecast can greatly improve results. Using fundamental data in a systematic way is perfectly consistent with other algorithmic approaches.

---

[7]Murray A. Ruggiero, Jr., "Fundamentals Pave Way to Predicting Interest Rates," *Futures* (September 1996).

Charting

I t is very likely that all trading systems began with a price chart, and we come back to a chart whenever we want a clear view of where the market is going. Nowhere can a picture be more valuable than in price forecasting. Elaborate theories and complex formulas may ultimately be successful, but the loss of perspective is easily corrected with a simple chart. We should remember the investor who, anxious after a long technical presentation by a research analyst, could only blurt out, "But is it going up or down?" Even with the most sophisticated market strategies, the past buy and sell signals should be seen on a chart. The appearance of an odd trade can save you a lot of aggravation and money.

Through the mid-1980s technical analysis was considered only as chart interpretation. In the equities industry, that perception is still strong. Most traders begin as chartists, and many return to it or use it even while using other methods. William L. Jiler, a great trader and founder of Commodity Research Bureau, wrote:

> *One of the most significant and intriguing concepts derived from intensive chart studies by this writer is that of characterization, or habit. Generally speaking, charts of the same commodity tend to have similar pattern sequences which may be different from those of another commodity. In other words, charts of one particular commodity may appear to have an identity or a character peculiar to that commodity. For example, cotton charts display many round tops and bottoms, and even a series of these constructions, which are seldom observed in soybeans and wheat. The examination of soybean charts over the years reveals that triangles are especially favored. Head and shoulders formations abound throughout the wheat charts. All commodities seem to favor certain behavior patterns.*[1]

In addition to Jiler's observation, the cattle futures market is recognized as also having the unusual occurrence of "V" bottoms. Until recently, both the silver and pork belly

---

[1] William L. Jiler, "How Charts Are Used in Commodity Price Forecasting," *Commodity Research Publications* (New York, 1977).

markets have tendencies to look very similar, with long periods of sideways movement and short-lived but violent *price shocks*, where prices leap rather than trend to a new level. The financial markets have equally unique personalities. The S&P traditionally makes new highs, then immediately falls back; it has fast, short-lived drops and slower, steadier gains. Currencies show intermediate trends bounded by noticeable major stopping levels while interest rates and bonds have long-term trends.

Charting remains the most popular and practical form for evaluating price movement, and numerous works have been written on methods of interpretation. This chapter will summarize some of the accepted approaches to charting and the trading rules normally associated with these patterns. Some conclusions are drawn as to what is most likely to work and why. The next chapter covers systems that are derived from these patterns and are designed to take advantage of behavioral patterns found in charts.

FINDING CONSISTENT PATTERNS

A price chart is often considered a representation of human behavior. The goal of any chart analyst is to find consistent, reliable, and logical patterns that can be used to predict price movement. In the classic approaches to charting, there are consolidations, trend channels, top-and-bottom formations, and a multitude of other patterns that are created by the repeated action of large groups of people in similar circumstances or with similar objectives. The most important of all the chart patterns is the *trendline*.

Only recently have computer programs been able to interpret chart patterns; and only one book, Bulkowski's *Encyclopedia of Chart Patterns*[2] has managed to show a comprehensive analysis of chart formations. In all fairness, there can be numerous valid interpretations of the same chart. In order to identify a chart price formation, it is first necessary to select the data frequency (for example, daily or weekly), then the starting date and a time horizon (long-term or short-term), before a chart interpretation can begin. Given the wide range of choices, it should be surprising that any two analysts see the same patterns at the same time.

Chart analyses, frequently published in magazines, may themselves be the cause of the repeated patterns. Novice speculators approach the problem with great enthusiasm and often some rigidity in an effort to follow to the rules. They will sell double and triple tops, buy breakouts, and generally do everything to propagate the survival of standard chart formations. Because of their following, it is wise to know the most popular techniques, if only as a defensive measure. Chapter 4 will review some of the attempts to turn these patterns into trading systems.

---

[2]Thomas N. Bulkowski, *Encyclopedia of Chart Patterns* (New York: John Wiley & Sons, 2000). Results of Bulkowski's studies are included in Chapter 4, in the section "A Study of Charting Patterns."

What Causes Chart Patterns?

Speculators have many habits, which, taken in large numbers, cause recognizable chart patterns. The typical screen trader (not on the exchange floor), or an investor placing his or her own orders, will usually choose an even number—for example, buy Microsoft at $26.00, rather than at $26.15. If even dollar values are not used, then 50¢ and 25¢ are the next most likely increments, in that order. And, as the share prices get higher, the increments get farther apart. With Berkshire Hathaway (BKA) trading at $125,000 per share, placing an order at a $10 increment would seem very precise. In futures trading, the same is true. There are far more orders placed in the S&P Index at 1310.00 than at 1306.50, or 10-year Treasury notes at $115\,{}^{16}\!/_{32}$ instead of $115\,{}^{19}\!/_{32}$.

The public is also said to always enter into the bull markets at the wrong time. When the television financial news, syndicated newspapers, and radio carry stories of dangerously low oil supplies, a new cancer treatment drug, or the devastation of the nation's wheat crop, the infrequent speculator enters in what W. D. Gann calls the *grand rush*, causing the final runaway move before the collapse or the final sell-off before the rally; this behavior is easily identifiable on a chart. Gann also talks of *lost motion*, the effect of momentum that carries prices slightly past its goal. Professional traders recognize that a fast, volatile price may move as much as 10% farther than its objective. A downward swing in the U.S. dollar/Japanese yen from par at 1.0000 to a support level of 0.8000 could overshoot the bottom by 0.0100 without being considered significant.

The behavioral aspects of prices appear rational. In the great bull markets, the repeated price patterns and divergence from chance movement are indications of the effects of mass psychology. The classic source of information on this topic is Mackay's *Extraordinary Popular Delusions and the Madness of Crowds* originally published in 1841.[3] In the preface to the 1852 edition the author says:

> *We find that whole communities suddenly fix their minds on one object, and go mad in its pursuit; that millions of people become simultaneously impressed with one delusion. . . .*

In 1975, sugar was being rationed in supermarkets at the highest price ever known, 50¢ per pound. The public was so concerned that there would not be enough at any price that they bought and horded as much as possible. This extreme case of public demand coincided with the price peak, and shortly afterwards the public found itself with an abundant supply of high-priced sugar in a rapidly declining market. The world stock markets are often the target of acts of mass psychology. While U.S. traders watched at a distance the collapse of the Japanese stock market from its heights of 38,957 at the end of December 1989 to its lows of 7,750 in 2003, a drop of 80%, they were able to experience their own *South Sea Bubble* when the NASDAQ 100 fell 83.5% from its highs of 4,816 in March 2000 to 795 in October 2002. And, while the subprime crisis has taken years to play out, the unparalleled drop in value of nearly all investments at the same time,

---

[3]Reprinted in 1995 by John Wiley & Sons.

September 2008, was clearly an act of investor panic. Prices seem to drop suddenly at the time when buyers are most confident, then start the long climb up again. It should not be difficult to understand why contrary thinking has developed a strong following.

Charting is a broad topic to be studied in detail; the chart paper itself and its scaling are sources of controversy. A standard bar chart (or line chart) representing highs and lows can be plotted for daily, weekly, or monthly intervals in order to smooth out the price movement over time. Bar charts have been drawn on semilog and exponential scales,[4] where the significance of greater volatility at higher price levels is put into proportion with the quieter movement in the low ranges by using percentage changes. Each variation gives the chartist a unique representation of price action. The shape of the chart box and its ratio of height/width will alter interpretations that are based on angles. Standard charting techniques may draw trendlines at 45° or 30° angles across the chart; therefore, expanding or compressing a chart on a screen will change the angles. This chapter uses traditional daily price charts and square boxes.

It may be a concern to today's chartist that the principles and rules that govern chart interpretation were based on the early stock market, using averages instead of individual stocks or futures contracts. This is discussed in the next section. For now, refer to Edwards and Magee, who removed this problem by stating that "anything whose market value is determined solely by the free interplay of supply and demand" will form the same graphic representation. They continued to say that the aims and psychology of speculators in either a stock or commodity environment would be essentially the same, that the effect of postwar government regulations have caused a "more orderly" market in which these same charting techniques can be used.[5]

WHAT CAUSES THE MAJOR PRICE MOVES AND TRENDS?

Prices can move higher for many months or even years, creating a *bull market.* They can also move down, creating a *bear market.* Although price moves can be as short as a few minutes or as long as decades (as happened with interest rates and gold), it is how each chartist defines a "trend" that is most important. Once recognized, the price trend forms a bias for trading decisions that can make the difference between success and failure. The long-term direction of prices is driven by four primary factors:

1. *Government policy.* When economic policy targets a growth rate of 4%, and the current growth rate is 1%, the Federal Reserve (the "Fed" or any central bank) lowers interest rates to encourage growth. Lowering rates stimulates business activity. The Fed raises interest rates and dampens economic activity to control inflation. Changing interest rates has a profound impact on the flow of investment money between countries, on international trade, on the value of currencies, and on business activity.

---

[4]R.W. Schabacker, *Stock Market Theory and Practice*, Forbes (New York, 1930), 595–600.

[5]Robert D. Edwards and John Magee, *Technical Analysis of Stock Trends* (Springfield, MA: John Magee, 1948), Chapter 16.

2. *International trade.* When the United States imports goods, it pays for it in dollars. That is the same as *selling* the dollar. It weakens the currency. A country that continually imports more than it exports increases its trade deficit and weakens its currency. A country that increases its exports strengthens its currency and its economy.

3. *Expectation.* If investors think that stock prices will rise, they buy, causing prices to rise. Expectations can lead an economic recovery although there is no statistical data to support a recovery. *Consumer confidence* is a good measure of how the public feels about spending. The economy is active when consumer confidence is high. A lack of public confidence following the subprime collapse dampened all economic activity and delayed the recovery for years.

4. *Supply and demand.* A shortage, or anticipated shortage, of any product causes its price to rise. An oversupply of a product results in declining prices. These trends develop as news makes the public aware of the situation. A shortage of a product that cannot be replaced causes a prolonged effect on its price, although the jump to a higher price may happen quickly.

THE BAR CHART AND ITS INTERPRETATION BY CHARLES DOW

The *bar chart*, also called the *line chart*, became known through the theories of Charles H. Dow, who expressed them in the editorials of the *Wall Street Journal.* Dow first formulated his ideas in 1897 when he created the stock averages in order to have a more consistent measure of price movement for stock groups. After Dow's death in 1902, William P. Hamilton succeeded him and continued the development of his work into the theory that is known today. Those who have used charts extensively and understand their weak and strong points might be interested in just how far our acceptance has come. In the 1920s, a New York newspaper was reported to have written:

> *One leading banker deplores the growing use of charts by professional stock traders and customers' men, who, he says, are causing unwarranted market declines by purely mechanical interpretation of a meaningless set of lines. It is impossible, he contends, to figure values by plotting prices actually based on supply and demand; but, he adds, if too many persons play with the same set of charts, they tend to create the very unbalanced supply and demand which upsets market trends. In his opinion, all charts should be confiscated, piled at the intersection of Broad and Wall and burned with much shouting and rejoicing.*[6]

This attitude seems remarkably similar to the comments about program trading that followed the stock market plunge in October 1987, where it was condemned as the cause of the crash. In 2011 we again had comments about high frequency trading "manipulating"

---

[6]Richard D. Wyckoff, *Stock Market Technique, Number One* (New York: Wyckoff, 1933), 105.

the markets, and in Europe they have banned short sales to stem volatility in the equity index markets. Of course, volatility continued to be high, but liquidity dropped. It's politics, not logic.

Charting has become an integral part of trading. The earliest authoritative works on chart analysis are long out of print, but the essential material has been recounted in newer publications. If, however, a copy should cross your path, read the original *Dow Theory* by Robert Rhea;[7] most of all, read Richard W. Schabacker's outstanding work *Stock Market Theory and Practice*, which is probably the basis for most subsequent texts on the use of the stock market for investment or speculation. The most available book that is both comprehensive and well written is *Technical Analysis of Stock Trends* by Edwards and Magee, now in its ninth edition.[8] It is focused on chart analysis with related management implications and a small section on commodities. For the reader who prefers concise information with few examples, the monograph by W. L. Jiler, *Forecasting Commodity Prices with Vertical Line Charts*, and a complementary piece, *Volume and Open Interest: A Key to Commodity Price Forecasting*, can still be found.[9] Two more recent publications that are widely read are John Murphy's *Technical Analysis of the Financial Markets* and Jack Schwager's *Schwager on Futures: Technical Analysis*, part of a two-volume set.

The Dow Theory

The Dow Theory[10] is still the foundation of chart interpretation and applies equally to stocks, financial markets, commodities, and the wide variety of investment vehicles used to trade them. It is part investor psychology supported by chart analysis. It is impressive that it has withstood the tests of more than 100 years. Charles Dow was the first to create an index of similar stocks—the Industrials and the Railroads, although today's components are very different from those in 1897. The purpose of the index was to smooth out erratic price movement and find consistency by combining less active stocks. Thin trading causes unreliable price patterns.

Dow's work can be viewed in two parts: his theory of price movement, and his method of implementation. Both are inseparable to its success. Dow determined that the stock market moved as the ocean, in three waves, called *primary*, *secondary*, and

---

[7]Arthur Sklarew, *Techniques of a Professional Chart Analyst* (Commodity Research Bureau, 1980).

[8]Robert D. Edwards and John Magee, *Technical Analysis of Stock Trends*, 9th ed. (Snowball Publishing, 2010).

[9]Two other works worth studying are Gerald Appel, *Winning Market Systems: 83 Ways to Beat the Market* (Great Neck, NY: Signalert, 1974); and Gerald Appel and Martin E. Zweig, *New Directions in Technical Analysis* (Great Neck, NY: Signalert, 1976).

[10]The rules of the Dow Theory in this section are based on a fine article by Ralph Acampora and Rosemarie Pavlick, "A Dow Theory Update," originally published in the *MTA Journal* (January 1978, reprinted in the *MTA Journal*, Fall–Winter 2001). Other parts of this section are drawn from Kaufman, *A Short Course in Technical Trading* (Hoboken, NJ: John Wiley & Sons, 2003).
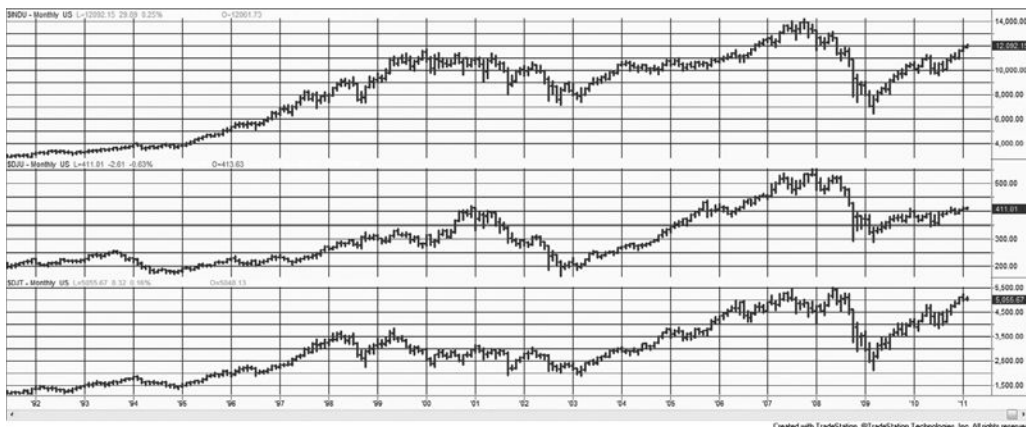
FIGURE 3.1 Dow Industrial, Utilities, and Transportation Indexes, 1991–2010. Dow originally created the industrial and railway averages to hide the large, erratic price moves caused by price manipulation and lack of liquidity. Dow Theory has been adapted to use the current versions of the major indexes, the Industrials (top panel), the Utilities (center panel), and the Transportation Index (bottom panel). Although these indexes represent different aspects of the economy, they have become highly correlated.

daily fluctuations. The major advances and declines, lasting for extended periods, were compared to the tides. These tides were subject to secondary reactions called *waves*, and the waves were comprised of *ripples*. Readers familiar with other charting methods will recognize these patterns as the foundation of Elliott Wave analysis. In 1897, Dow published two sets of averages in the *Wall Street Journal*, the *Industrials* and the *Railroads* in order to advance his ideas. These are now the *Dow Jones Industrial Average* and the *Transportation Index*. Figure 3.1 shows more than 20 years of history for the three most important averages the Industrials, the Transportation, and the Utilities.

The Basic Tenets of the Dow Theory

There are six fundamental principles of the Dow Theory that fully explain its operation.

1. The Averages Discount Everything (except "acts of God")

At the turn of the twentieth century there was considerably less liquidity and regulation in the market; therefore, manipulation was common. By creating averages, Dow could re-duce the frequency of "unusual" moves in a single stock, that is, those moves that seemed unreasonably large or out of character with the rest of the market. Dow's Industrials average the share value of 30 companies (adjusted for splits); therefore, an odd move in one of those prices would only be $\frac{1}{30}$ of the total, reducing its importance so that it would not distort the results. The average also represented far greater combined liquidity than a single stock. The only large moves that would appear on a chart of the average price were price shocks, or "acts of God."

2. Classifications of Trends

There are three classifications of trends: primary trends, secondary swings, and minor day-to-day fluctuations. The primary trend, also called the *wave*, is the trend on a grand scale. When there is a wave of rising prices we have a *bull market*; when prices are declining there is a *bear market*. A wave is a major move over an extended period of time, generally measured in years. A clear bull market can be seen in the previous Dow charts (Figure 3.1) throughout all of the 1990s ending at the beginning of 2000, and again from 2003 through mid-2007.

Bull and Bear Market Formation (for Monthly or Weekly Prices)

The beginning of a bull or bear market is determined using a *breakout signal*, shown in Figure 3.2, based on large swings in the index value (a complete explanation of breakout signals can be found in Chapter 5). The *bull market signal* occurs at the point where prices confirm the uptrend by moving above the high of the previous rally. The *bear market signal* occurs on a break below the low of the previous decline.

It is commonly accepted that a bull or bear market begins when prices reverse 20% from their lows or highs. In order to get an upwards breakout signal needed for a new bull market, we want to look at support and resistance levels (the previous intermediate high and low prices) separated by approximately a 10% price move based on the index value. This type of signal is called *swing trading.* At the top of Figure 3.2 the horizontal broken line should occur at about 20% below the absolute price highs, and the second peak should be approximately 10% higher than the previous swing low.

It is interesting to note that both bull and bear markets start with a price reversal of 20%. But 20% from the highs can be much greater than 20% from the lows. For example, in the sell-off in September 2008, the S&P was measured from its high of about 14,000 in late
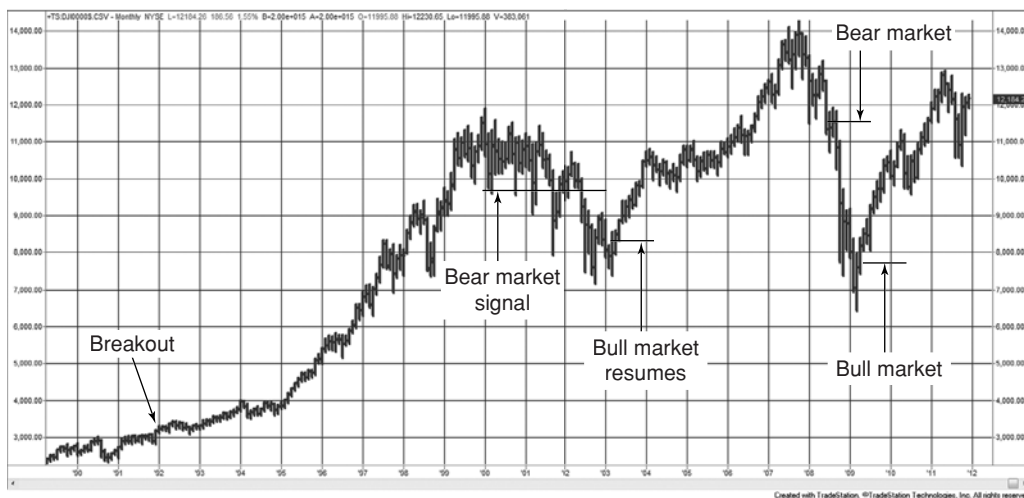
FIGURE 3.2   Bull and bear market signals are traditional breakout signals, but on a larger scale.

2007. A decline to 11,200, or 2,800 points, triggered the bear market. In the first quarter of 2009, the S&P reached its lows of about 6,500. A new bull market began at 7,800, a rally of only 1,300 points. Thus the number of points needed to "officially" start a bull market was only 46% of the bear market trigger, showing a significant bias toward bull markets.

Bull and Bear Market Phases

In Dow Theory, the primary trends develop in three distinct phases, each characterized by investor action. These phases can be seen in the NASDAQ bull market of the late 1990s and the subsequent bear market (Figure 3.3).

The Bull Market

> *Phase 1: Accumulation.* Cautious investors select only the safest and best-valued stocks to buy. They limit purchases to deeply discounted stocks at depressed price levels and consider only primary services and industries, most often buying utilities and high yielding stocks.
>
> *Phase 2: Increasing volume.* Greater investor participation causes increasing volume, rising prices, and an improving economic picture. A broader range of investors enters the market convinced that the market has seen its lowest prices. Secondary stocks become popular.
>
> *Phase 3: Final explosive move.* Excessive speculation and an elated general population result in a final explosive move. Everyone is talking about the stock market;

FIGURE 3.3 NASDAQ from April 1998 through June 2002. A clear example of a bull and bear market with a classic pattern of volume.

people who have never considered investing directly now enter the market. The public is convinced that profits will continue and buying becomes indiscriminate. Investors borrow to buy stocks. Value is unimportant because prices keep rising. Earnings and dividends are ignored.

The Bear Market

Phase 1: Distribution. Professionals begin selling while the public is in the final stages of buying. Stocks are distributed from stronger to weaker hands. The change of ownership is facilitated by less experienced investors who enter the bull market too late and pay what turn out to be unreasonably high prices.

Phase 2: Panic. Prices decline faster than at any time during the bull market and fail to rally. The news constantly talks about the end of the bull market. The public sees an urgency to liquidate. Investors who borrowed money to invest late in the bull market, trading on margin or leverage, now speed up the process. Some are forced to liquidate because their portfolio value has dropped below the critical point. The divesting of stocks takes on a sense of panic.

Phase 3: Lack of buying interest. The final phase in the sustained erosion of prices results from the lack of buying by the public. After taking losses, investors are not interested in buying even the strongest companies at extremely undervalued prices. All news is viewed as negative. Pessimism prevails. It is the summer of 2002.

Schabacker's Rules

Schabacker also had a simple guideline to identify the end of both a bull and a bear market.[11]

End of a Bull Market
1. Trading volume increases sharply.
2. Popular stocks advance significantly while some other companies collapse.
3. Interest rates are high.
4. Stocks become a popular topic of conversation.
5. Warnings about an overheated stock market appear on the news.

End of a Bear Market
1. Trading volume is low.
2. Commodity prices have declined.
3. Interest rates have declined.
4. Corporate earnings are low.
5. Stock prices have been steadily declining and bad news is everywhere.

---

[11] Adapted from James Maccaro, "The Early Chartists: Schabacker, Edwards, Magee," *Technical Analysis of Stocks & Commodities* (November 2002).

Secondary Trends (Secondary Reactions Using Weekly or Daily Prices)

Secondary reactions are also called *corrections* or *recoveries* and can be identified using smaller swing values. Corrections in bull markets are attributed to the prudent investor taking profits. This profit phase can have an erratic start but is considered complete when prices rise above the previous secondary rally. The bull market is back in force when a new high occurs (see Figure 3.4), the point where a trader can enter a new long position. *Lines* may be substituted for secondary movements. In Dow Theory, a *line* is a sideways movement lasting from two to three weeks to months, trading in about a 5% range.

Characteristics of a Secondary Reaction
- There are a number of clear downswings.
- The movement is more rapid in the reversal (down during a bull market) than in the primary move.
- The reactions last from three weeks to three months.
- If the volume during the price drop is equal to or greater than the volume just prior to the decline, then a bear market is likely. If volume declines during the drop, then a reaction is confirmed.
- The atmosphere surrounding the decline is important. If there is a lot of speculation, then a bear market may develop.

Minor Trends (Using Daily Prices)

In Dow Theory, minor trends are the only trends that can be manipulated. They are usually under six days in duration. Because they are considered market noise, not affecting the major price direction, they are seen as frequent up and down movements.
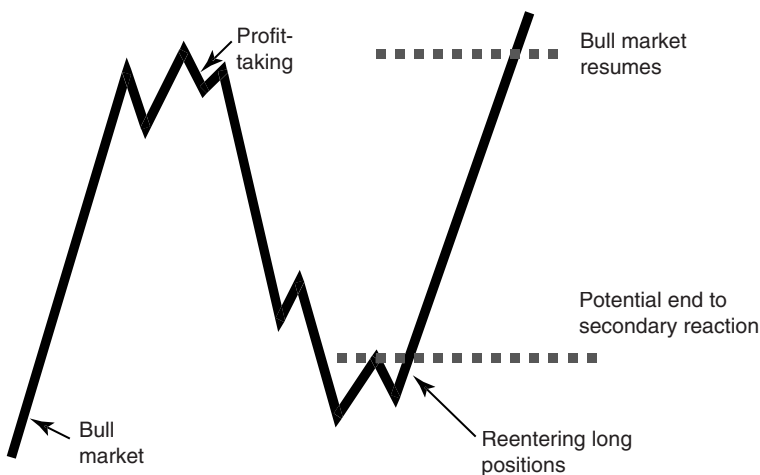
FIGURE 3.4 Secondary trends and reactions. A reaction is a smaller swing in prices that ends when a new high reinstates the bull market.

3. The Principle of Confirmation

For a bull or bear market to exist, two of the three major averages (the Industrials, the Transportation, and the Utilities) must confirm the direction. When first created, the Dow Theory required the confirmation on only the Utilities and the Railroads. Although much has changed since Dow devised this rule, the purpose is to assure that the bull or bear market is a widespread economic phenomenon and not a narrower industry-related event.

4. Volume Goes with the Trend

Volume confirms the price move. Volume must increase as the trend develops, whether it is a bull or bear market. It is greatest at the peak of a bull market or during the panic phase of a bear market.

5. Only Closing Prices Are Used

Dow had a strong belief that the closing price each day was the most important price. It was the point of evening-up. Not only do day traders liquidate all of their positions before the close of trading, reversing their earlier impact, but many investors and hedge funds execute at the close. Although liquidity was a problem during Dow's time, even actively traded stocks in today's market show increased price swings when a larger order is executed during a quiet period. There is always high volume at the close of trading, when investors with short and long time frames come together to decide the fair price.

Some traders believe that there is no closing price anymore, given the access to 24-hour trading; however, that is not yet true. Every market has a settlement price. This is usually at the end of the primary trading session (previously the *pit* or *open outcry session*). The settlement price is necessary to reconcile all accounts, post profits and losses, and trigger needed margin calls. Banks could not operate without an official closing time and settlement price.

6. The Trend Persists

A trend should be assumed to continue in effect until its reversal has been signaled. This rule forms the basis of all trend-following principles. It considers the trend as a long-term price move, and positions are entered only in the trend direction. The Dow Theory does not express expectations of how long a trend will continue. It simply follows the trend until a signal occurs that indicates a change of direction.

Interpreting Today's S&P Using Dow Theory

After 110 years, can the Dow Theory correctly interpret the major market index, the S&P? Figure 3.5 shows the S&P 500, using continuous, back-adjusted futures prices, from 1994 through the middle of 2003. The sustained bull market that began in 1987, or possibly 1984, peaks near the end of the first quarter of 2002. There is a steady increase in volume, as Dow had foreseen, although volume does not peak at the top of the market—it starts to decline noticeably about three months before the top. We will see in the study of volume that volume spikes occur at extremes, but a longer-term volume confirmation is very important. Declining volume at the beginning of 2003 signals a divergence in sentiment

FIGURE 3.5 Dow Theory applied to the S&P. Most of Dow's principles apply to the current marketplace, but some experience and interpretation is necessary.

that foretells the end of the bull market. Volatility increases as prices move towards the end of the uptrend, another predictable pattern. The price move from 1994 through the peak in 2002 shows both Phase 2 and Phase 3 of the bull market.

The price decline in the third quarter of 1998 addresses the issue: Are there exceptions to the 20% rule that changes a bull market to a bear market? A 20% drop from a high of 1400 is 1120, very close to the point where prices stopped their decline and reversed. Dow never used the number 20%, and analysts would claim that, because of the speed of the decline and the quick recovery, this was not a bear market signal. Some of these decisions require judgment, some experience, and just a little bit of hindsight. Realistically, we cannot expect every Dow signal to always be correct, just as we cannot expect to be profitable on every trade. Long-term success is the real goal.

Transition from Bull to Bear in the S&P

Looking again for a 20% reversal from the S&P highs of 1675, we target the price of 1340. This time, volume has declined into the highs and continues to decline quickly. From the second quarter of 2000 through the first quarter of 2001 prices fall sharply, giving back the gains from mid-1997, nearly three years. When prices break below 1300 they confirm the previous low at the end of 2000, making it clear that a bear market is underway.

During the subsequent decline, prices attempted to rally. There are four cases of a sharp "V" bottom followed by a significant move higher. After the low at 940 at the end of September 2001, prices move to about 1180, above the 20% reversal of 1128. However, after the first reversal to 1075 prices fail to move back above the highs, finally breaking below 1180 and continuing on to make new lows. Although the recovery exceeded 20%,

the lack of a confirming breakout can be interpreted as a bull market failure. Not every pattern falls neatly into a rule.

We come to the last year of the S&P chart, where prices have resisted going below 850, and now appear to be moving above the level of 970 and about to confirm a bullish breakout. Is it the end of the bear market? Volume was the highest at the two lowest price spikes, and then declined. Many stocks are undervalued, according to experts, yet those same experts see no reason for the market to rally further because the recent rise has already reflected reasonable expectations for profits and growth in the next year. Who would be correct, Charles Dow or the talking heads of the financial news networks? It was Dow.

Dow Theory and Futures Markets

The principles of the Dow Theory are simple to understand. Major price moves are most important when they are confirmed by volume. They follow a pattern created by investor action that seems to be universal when seen from a distance. In order to implement his theory, Dow created an index that minimized the erratic moves in individual stocks due to lack of liquidity and price manipulation.

The primary features of the Dow Theory should hold for any highly liquid, actively traded market. This applies to index futures and most financial futures markets, as well as foreign exchange, which have enormous volume and reflect major economic trends. Because of the variety of products traded as futures and ETFs, an investor may be able to apply Dow's principle of confirmation using any two related financial markets, such as the S&P Index, 10-year Treasury notes, or the U.S. dollar index, in the same way that the Industrials, Utilities, and Transportation indexes were used for stocks. A strong economic trend often begins with interest rate policy and has a direct impact on the value of the currency, and a secondary effect on the stock market. Stock prices can be stimulated by lower rates or dampened by raising rates; therefore, confirmation from these three sectors is reasonable. When trading in futures, the nearby contract (the one closest to delivery) is most often used; however, the total volume of all futures contracts traded for each market must be used rather than volume for a single contract.

CHART FORMATIONS

While Dow Theory is a macro view of price movement, more often chart analysis deals with much shorter time periods. Most traders hold positions from a few days to a few weeks; however, they apply the same patterns to both shorter or longer intervals.

Chart analysis uses straight lines and geometric formations on price charts. It analyzes volume only in the most general terms of advancing and declining phases. Chart patterns can be classified into the broad groups of:

- Trendlines and channels
- One-day patterns

Charting   **93**

- Continuation patterns
- Accumulation and distribution (tops and bottoms)
- Retracements
- Other patterns

Of these, the most important is the *trendline*.

The Trend in Retrospect

It is easier to see the trend on a chart after it has occurred. Trying to identify the trend as it is developing is much more difficult. The monthly chart in Figure 3.6 shows a sustained uptrend trend, but there is a slowing of that trend toward the end. Will the upward trend continue? Will prices begin a downward trend? Will they move sideways? The purpose of charting is to apply tools that provide the best chance of identifying the future direction of prices. If wrong, these tools also control the size of the loss.

The time interval is a key element when identifying a trend. Weekly and monthly charts show the major trends more clearly than daily charts. Longer-term charts remove much of the noise that interferes with seeing the bigger picture. Many chartists start by evaluating a weekly or monthly chart, then apply the lines and values developed on those charts to a daily chart. The weekly chart provides direction or biases the direction of trades while the daily chart, or even a 15-minute chart, is used for timing entries and exits. Further discussion of this can be found in Chapter 19.

FIGURE 3.6 The trend is easier to see after it has occurred. While the upwards trend is clear, are prices going to continue higher, or is this the end of the trend?

TRENDLINES

The trendline determines the current direction of price movement, and often identifies the specific point at which that direction will change. The trendline is the most popular and recognized tool of chart analysis. Most analysts will agree that *the trend is your friend*; that is, it is always safer to take a position in the direction of the trend.

- An *upwards trendline* is drawn across the lowest prices in a rising market.
- A *downwards trendline* is drawn across the highest prices in a declining market.

Figure 3.7 shows a classic downwards trendline, *A*, drawn on a chart of Intel. It connects the highest price of $22 with price peaks at 18.00, 16.75, and 16.15 before ending at 15.50. When prices move through the trendline heading higher, the downtrend has been *penetrated*. This may end the downtrend or cause a new downtrend line to be drawn. In this case it was the end of the downtrend.

Redrawing Trendlines

Most trendlines are not as long-lived or clear as the downtrend in Intel, which was drawn after the fact. Instead, we will treat the uptrend as it develops. The first uptrend line, *B*, is drawn when the first reversal shows a second low point. The upwards trendline *B* is drawn across the lows of points 1 and 2. Although prices do not decline through trendline *B*, rising prices pull back to points 3 and 4, well above the trendline.

FIGURE 3.7   Upwards and downwards trendlines applied to Intel, November 2002 through May 2003.

At that point, we choose to redraw the upwards trendline connecting point 2 with 3 and 4, forming what appears to be a stronger trendline. Trendlines are considered more important when they touch more points. However, prices move up quickly, and we decide to redraw the trendline connecting points 4 and 5. It is very common to redraw trendlines as price patterns develop. Care must be taken to draw the lines in a way that touches the most points, although some chart analysts would draw a line that connects points 1 and 5, crossing through points 2, 3, and 4, because the final picture seems to represent the dominant upwards price pattern. This can be seen as the broken line in Figure 3.7.

Support and Resistance Lines

Price movement creates patterns that reflect the combined perception that all investors have of the current economic situation. Trends result from confidence or concern about the health of business or the supply and demand of a product. When there is no dominant opinion, prices move sideways in a price range determined by current volatility levels—sometimes wide, sometimes narrow. Because there are always buyers and sellers, prices do not stand still. Investment funds continue to add and withdraw money from the market.

Periods of uncertainty form a sideways price pattern. The top of this pattern is called the *resistance level*, and the bottom is the *support level*. Once established, the support and resistance levels become key to identifying whether a trend is still in force.

A *horizontal support line* is drawn horizontally to the right of the lowest price in a sideways pattern. It is best when drawn through two or more points and may cross above the lowest price if it makes the pattern clear. It represents a firm price level that has withheld market penetration (or allowed minor penetration). It may be the most significant of all chart lines. In the chart of gold futures prices (Figure 3.8), the support line is drawn across the bottom of a sideways period, beginning at the first low price on the left but crossing slightly above the next lowest point. The support line could have been drawn at \$280.50 to include the first cluster of low prices and crossing above the lows bars but representing a clear support level.

A *horizontal resistance line* serves the same purpose as the support line and is drawn across the highest highs of the sideways interval. It represents the price that has resisted upwards movement. Resistance lines are not normally as clear as support lines because they are associated with higher volatility and erratic price movement. In Figure 3.8 there are two choices for the horizontal resistance line. The most common selection would be the line that begins at top 1 and crosses below the high of top 2. In the same spirit as the support line, a resistance line could have been drawn much lower, beginning at top 3 and crossing above a cluster of highs while penetrating through the bars with tops 1 and 2.[12]

---

[12]In Carol Oster, "Support for Resistance: Technical Analysis and Intraday Exchange Rates," *FRBNY Economic Policy Review* (July 2000), the author shows that support and resistance levels specified by six trading firms over three years were successful in predicting intraday price interruptions. In addition, these levels were valid for about five days after they were noted.
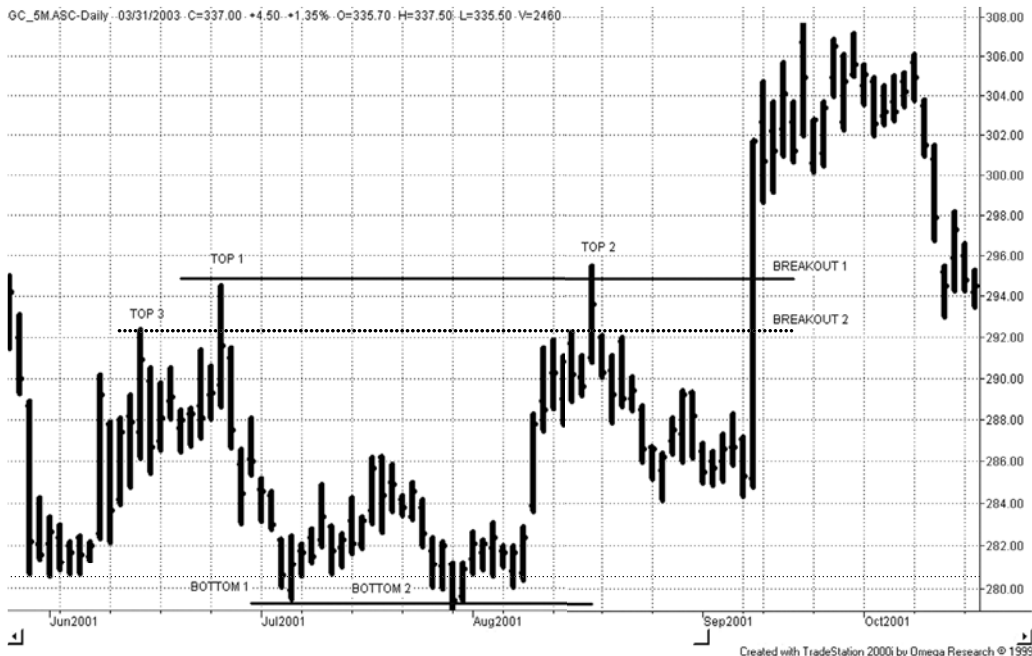
FIGURE 3.8 Horizontal support and resistance lines shown on gold futures prices.

Note the Position of the Closing Price of the Bar

A price bar that has the high price penetrating upwards through resistance but closes lower is considered a *failed breakout* and confirms the sideways pattern. The same is true for a failed penetration of the support level. You may choose to raise the resistance line to the high of that failed bar, but most chartists ignore it, keeping the resistance line at its original position. Then we can expect to see a number of high prices penetrate through the resistance lines as shown by the *breakout 2* line in Figure 3.8.

Resistance Becomes Support, and Support Becomes Resistance

Horizontal support and resistance lines are strong indicators of change. If prices are moving sideways because investors are unsure of direction, then a move through either support or resistance is usually associated with new information that causes investors to act. Whatever the cause, the market interprets this as a new event. Having moved out of the sideways pattern, prices have a tendency to remain above resistance to confirm the change. If prices have moved higher, then the resistance line becomes a support line. If prices fall below the resistance line, the price move is considered a failed breakout. In the right part of Figure 3.8, prices break out above the resistance levels and then come back to test those levels. In this example, prices seem to confirm that the *breakout 2* line was the more realistic resistance line.

A Trendline Is a Support or Resistance Line

The angled trendlines in Figure 3.7 are also called support and resistance lines. An upwards trendline, drawn across the lows, is a *bullish support line* because it defines the

lowest price allowed in order to maintain the upwards trend. The downward trendline, drawn across the highs, is a *bearish resistance line*. These angled trendlines are most reliable when used to identify major price trends. Horizontal lines work well for shorter time frames.

Back-Adjusted Data

All traders use online services to display charts. They can draw support and resistance lines using various tools supplied by the service and can convert daily data to weekly or monthly with a single click. When looking at prices that go back many years, the analyst must be sure that the older data is not *back-adjusted* in any way. For example, futures trade in contracts of limited maturity, and are most liquid during the last few months before expiration. Long-term charts put contracts together by back-adjusting the prices, so that the older data does not give the actual price at that time, but is altered by accumulated roll difference. Using those older prices as a guide for support or resistance doesn't make any sense.

This same problem exists for stocks that have split. The old price that you see on the chart may not be the actual price traded at that time. Floor traders are good at remembering the last major high or low and will trade against those prices, or buy and sell breakouts. When plotting the support and resistance lines, look for the data option that creates a history of prices without back-adjusting.

Rules for Trading Using Trendlines

The simplest formations to recognize are the most commonly used and most important: horizontal support and resistance lines, bullish and bearish support and resistance lines, and channels created using those lines. Proper use of these basic lines is essential for identifying the overall direction of the market and understanding the patterns formed as prices move from one level to another. Many traders will generate their buy and sell orders directly from their chart analysis. Other, more computer-oriented analysts have automated the more important trendlines, particularly horizontal support and resistance, which has become the basic breakout system. Major chart patterns create the underlying profitability of chart trading; the more complex formations, as we will discuss further, may enhance good performance but rarely compensate for losses resulting from being on the wrong side of the trend.

Once the support and resistance lines have been drawn, a price penetration of those lines creates the basic trend signal (Figure 3.9). The bullish support line defines the upward trend, and the bearish resistance line denotes the downward one. For long-term charts and major trends, this is often sufficient. Some traders add the additional rule that once the price has penetrated a trendline, it must remain penetrated for some time period in order to confirm the new trend. Most false penetrations correct quickly.

Confirming the New Trend Direction

In actual trading, the price crossing the trendline is not as clean as in Figure 3.8. Most often prices that have been moving higher will cross below the trendline, then recross
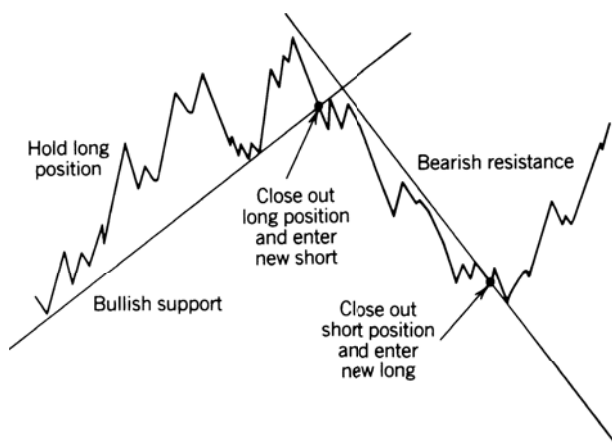
FIGURE 3.9   Basic sell and buy signals using trendlines.

moving higher, then move lower again. The trendline is an important turning point, and there may be indecision that is reflected in a sideways price movement before prices reestablish a trend. To deal with this situation, traders may:

- *Wait a set time period* to confirm that prices remain on the new side of the trendline.
- *Wait for a reversal* after the penetration, then enter a trade in the new direction even if the reversal crosses the trendline again.
- *Create a small safety zone* (called a *band* or *channel*) around the trendline and enter the new trade if prices move through the trendline and through the safety zone.

Each of these techniques requires a delay before entering. A delay normally benefits the trader by giving a better entry price; however, if prices fall quickly through an up-wards trendline and do not reverse or slow down, then any delay will result in a much worse entry price. Unfortunately, most of the biggest profits result from breakouts that never pull back. Catching only one of these breakouts can compensate for all the small losses due to false signals. Many professional traders wait for a better entry price. They may be steady winners, but they do not often profit from the biggest moves.

Trading Rules for Horizontal Support and Resistance Levels

As with angled trendlines, horizontal support and resistance lines show clear points for buying and selling. Also similar to angled trendlines, the horizontal lines become increasingly important when longer time intervals and more points are used to form the lines. The technique for entering trades using horizontal lines is similar to that using angled trendlines; however, the maximum risk of the trade is clearly defined.

- *Buy* when prices move above the horizontal resistance line
- *Sell* when prices move below the horizontal support line
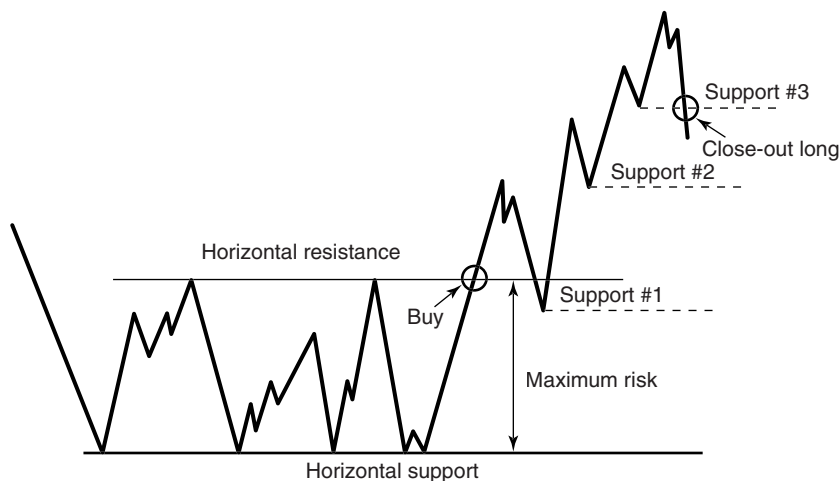
FIGURE 3.10   Trading rules for horizontal support and resistance lines.

Once a long position has been entered, it is not closed out until prices move below the support line. The maximum risk of the trade is the difference between the support and resistance lines. As prices move higher, each swing reversal forms a low from which a new horizontal support line is drawn. After the initial entry, single points are most often used to create the horizontal support and raise the level at which the trade will be closed out. Figure 3.10 shows the pattern of horizontal support and resistance lines as the trade develops. For a swing low to form, prices must reverse by more than some threshold number of points or percentage. Not every small reversal qualifies as a swing low.

Note that the first pullback in Figure 3.10 shows prices crossing below the original resistance line. This is a common occurrence, but the original line no longer holds the importance it had before it was broken. While it should provide support for the pullback (a resistance, once broken, becomes a support), it is more important to record the bottom of the new pullback as the support level. These new support levels need only one price point. After the third support level is drawn, prices rally but then fall back through the third level, at which point the long position is closed out. A short position, if any, is not entered until a new sideways price pattern is established and horizontal support and resistance lines can be drawn across more than one point.

Identifying Direction from Consolidation Patterns
It is said that markets move sideways about 80% of the time, which means that sustained directional breakouts do not occur often, or that most breakouts are false and fail to identify a new market direction. Classic accumulation and distribution formations, which occur at long-term lows and highs, attempt to find evolving changes in market sentiment. Because these formations occur only at extremes, and may extend for a long time, they represent the most obvious consolidation of price movement. Even a rounded or saucer-shaped bottom may have a number of false starts; it may seem to turn up in a

uniform pattern, then fall back and begin another slow move up. In the long run, the pattern looks as if it is a somewhat irregular, extended rounded bottom; however, using this pattern to enter a trade in a timely fashion can be disappointing. It is easier to average *in*, where smaller positions are entered at fixed intervals as long as the developing formation remains intact.

Most other consolidation formations are best viewed in the same way as a simple horizontal sideways pattern, bounded above by a resistance line and below by a support line. If this pattern occurs at reasonably low prices, we can eventually expect a breakout upwards when the fundamentals change. Occasionally, prices seem to become less volatile within the sideways pattern, and chartists take this opportunity to redefine the support and resistance levels so that they are narrower. Breakouts based on these more sensitive lines tend to be less reliable because they represent a temporary quiet period inside the larger, normal level of market noise; however, there are two distinct camps, one that believes that breakouts are more reliable after a period of low volatility and the other that prefers breakouts associated with high volatility. Some of the situations are discussed in Chapter 20 under "VIX" and "Volatility System."

Creating a Channel with Trendlines

A *channel* is formed by a trendline and another line drawn parallel to the trendline enclosing a sustained price move. The purpose of the channel is to define the volatility of the price move and establish reasonable entry and exit points. Up to now, the trendline has only been used to identify the major price direction. A long position is entered when the price crosses a downward trendline moving higher. The trade is held until the price moves below the upwards trendline. However, it is more common to have a series of shorter trades. While the biggest profits come from holding one position throughout a sustained trend, a series of shorter trades each has far less risk and is preferred by the active trader. Be aware that trendlines using very little data are essentially analyzing noise and have limited value.

Before a channel can be formed, the bullish or bearish trendline must be drawn. A clear uptrend line requires at least two, and preferably three or more major low points on the chart, as shown in Figure 3.11, where points 1, 2, and 3 are used. These points do not have to fall exactly on the line. Once the trendline is drawn, the highest high, point *B*, can be used to draw another line parallel to the upwards trendline. The area in between the two parallel lines is the *channel*.

In theory, trading a channel is a simple process. We buy as prices approach the support line (in this case the upwards trendline), and we sell as prices near the resistance line. These buy and sell zones should be approximately the bottom and top 20% of the channel. Because the channel line is used to determine price targets, you might choose to draw the broken line across point *A*. The use of point *A* creates a channel that is narrower than the one formed using the higher point *B* and recognizes the variability of price movement. This allows you to take profits sooner.

If prices continue through the lower trendline after a long position has been set, the trade is exited. The trend direction has changed, and a new bearish resistance line,
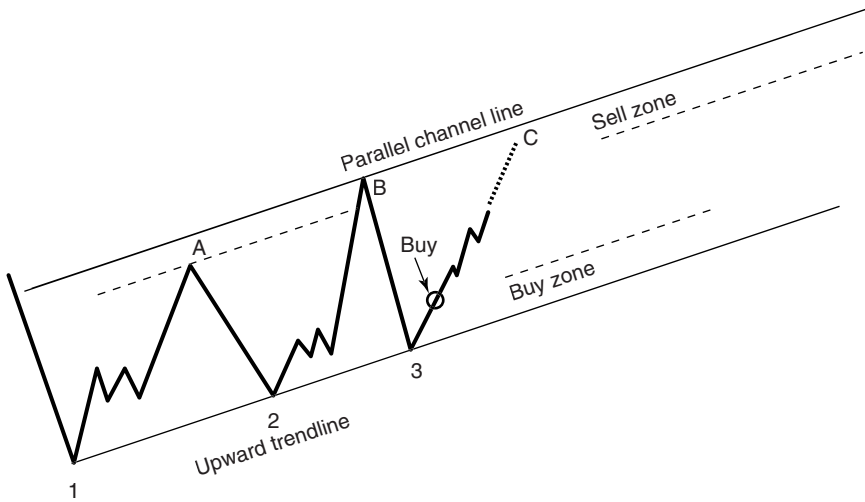
FIGURE 3.11 Trading a price channel. Once the channel has been drawn, buying is done near the support line and selling near the resistance line.

the downward trendline, needs to be drawn using points *B* and *C*, shown in Figure 3.12. Once the first pull-back occurs leaving a low at point 4, a parallel line is drawn crossing point 4, forming the downward channel. In a downward trending channel, it is best to sell short in the upper zone and cover the short in the lower zone. Buying in the lower zone is not recommended; trades are safest when they are entered in the direction of the trend.

When the support and resistance lines are relatively horizontal, or sideways, the channel is called a *trading range*. There is no directional bias in a trading range; therefore,
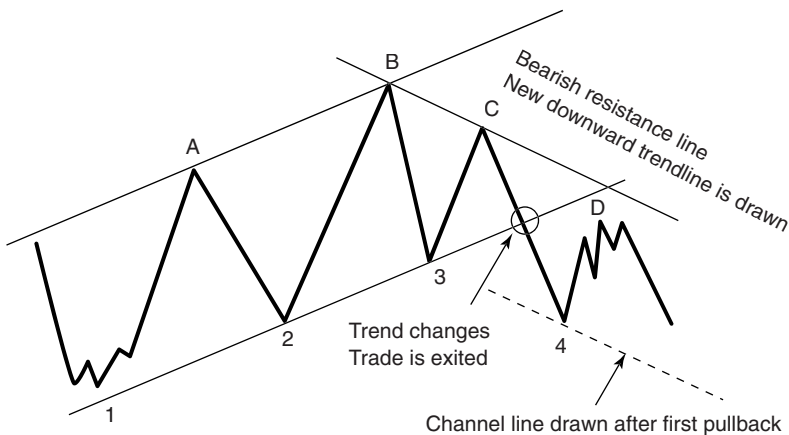
FIGURE 3.12 Turning from an upward to a downward channel. Trades are always entered in the direction of the trend.

you can enter new long positions in the support zone and enter new shorts in the sell zone. In both cases, penetration of either the support or resistance lines forces liquidation of the trade and establishes a new trend direction.

ONE-DAY PATTERNS

The easiest of all chart patterns to recognize occur in one day. They include gaps, spikes, island reversals, reversal days, inside days, outside days, wide-ranging days, and, to a lesser extent, thrust days. Some of these patterns are important at the moment they occur, and others must be confirmed by other factors.

Gaps

Price *gaps* occur when important news influences the market at a time when the exchange is closed. Orders accumulate to be executed on the next open. An *upwards gap* exists when the low of the current day is higher than the high of the previous day. If all trading were 24 hours, then we would see a fast, volatile move, but not a gap. For example, there are three popular ways to trade the S&P 500 index,

1. Futures, traded in the Chicago Mercantile Exchange pit from 8:30 A.M. to 3:15 P.M. (Central time)
2. Spyders (SPY) on the AMEX during the same hours
3. The electronic S&P mini-futures contract which trades nonstop from Sunday evening at 6 P.M. until Friday afternoon at 3:15 P.M. on the Globex platform

Gaps only exist when using the primary trading session (the *pit* session). Most afterhours trading is on light volume and may be ignored for charting purposes. An exception is in Europe where they have an extended session from the original 4 P.M. close (European time) to 10 P.M. or 10:30 P.M. to allow trading at the same time the U.S. markets are open. The markets then close and reopen when the normal European business day begins. For European markets it is best to use the combined sessions that start at about 9 A.M. and continue until about 10 P.M.

Economic reports are released by the U.S. government at 7:30 A.M. (Central time); therefore, they occur before the S&P pit trading and the SPDRs begin trading, but during the electronic *e*mini S&P session. There is no gap in electronic trading, but the other markets will open sharply higher or lower to adjust to the current *e*mini price. When the financial news shows give the expected open of the stock market, they are using the difference between the previous NYSE closing price and the current price of the electronic session. This creates frequent opening gaps in those markets.

Gaps can also occur because of a large cluster of orders placed at the point where the stock or futures market penetrates support or resistance. It is possible to have a gap during the trading session, immediately following bullish economic report, or

concurrent with an anticipated news release of consequence, when there are a large number of buyers and few sellers. There is also the rarer case of an event shock such as September 11, 2001.

In charting, gaps are interpreted differently based on where they occur in the current price pattern. In some cases, a gap signals a continued move and in other situations it is expected to be the end of a price move. The four primary gap formations are shown on a chart of Amazon.com in Figure 3.13. They are:

1. The *common gap*, which appears as a space on a chart and has no particular attributes—that is, it does not occur at a point associated with any particular significance. A common gap appears in May 1999 during a downward move.
2. A *breakaway gap* occurs at a point of clear resistance or support. It occurs when there are a large number of buy orders just above a major resistance line, or sell orders below a support line. Most often this is seen after a prolonged period of sideways price movement when most chartists can draw the same horizontal support and resistance lines. The clearer the formation, and the longer the sideways period, the more likely there will be a large breakaway gap. The term *breakaway* requires some hindsight because it is applied only when the gap is followed by a sustained price move.

    There are two breakaway gaps in Figure 3.13, the first shortly after prices make a new high, the second in the middle of the chart when prices break upwards through

FIGURE 3.13 Price gaps shown on a chart of Amazon.com.

a steep downward trendline, and the last near the right of the chart when prices gap through a clear downwards trendline.

   In order to trade a gap, a position must be entered in advance of the gap, as prices approach the support or resistance level. Once a long position is set and prices *gap up* you gain *free exposure*, which is the profit caused by the gap or by a fast market move in your favor. If prices do not gap up, they most often drift lower. The position can be exited with a small loss and reentered later.

3. An *exhaustion gap* occurs at the end of a sustained and volatile price move and confirms the reversal. Exhaustion gaps usually occur on the day after the highest price of the upwards move; however, in the Amazon chart, it is one day later. Because it signifies a clustering of orders anxious to exit the long side, it has all the signs associated with an exhaustion gap.

4. A *runaway gap* occurs at different points during a clear trend and confirms the trend. It does not appear to have any practical use because the trend can stop and reverse just after a runaway gap and it will be renamed an *island top*, or some other formation. When holding a long position, an upwards runaway gap quickly adds profits, but also signifies extreme risk.

Gaps can also be a hindrance to trading. A long position held when a downward breakaway gap occurs guarantees that any stop-loss order is executed far away from the order price. If the upwards breakaway gap occurs on light volume, it may be a false breakout. If a short is held, and if you are lucky, prices will fall back to the breakout level and then continue lower. If unlucky, you will be executed at the high of the move. In the final analysis, if the gap breakout represents a major change, a trade should be entered immediately *at the market*. The poor executions will be offset by the one time when prices move quickly and no pullback occurs. A breakaway gap on high volume is usually indicative of a strong move and a sustained change.

Filling the Gap

Tradition states that prices will retrace to *fill the gap* that occurred sometime earlier. Naturally, given enough time, prices will return to most levels; therefore, nearly all gaps will eventually be filled. The most important gaps are not filled for some time.

   The gap represents an important point at which prices move out of their previous pattern and begin a new phase. The breakaway gap will often occur just above the previous normal, or established, price level. With commodities, once the short-term demand imbalance has passed, prices should return to near-normal (perhaps slightly above the old prices given inflation), but also slightly below the gap. When a stock price gaps higher based on earnings, a new product announcement, or a rumor of an acquisition, the price may not return to the previous level.

   A computer study of opening gaps, and a program that performs the study, can be found in Chapter 15. It shows the probabilities of subsequent price moves following gaps, based on different-sized gaps, in a wide selection of markets. Bulkowski has a large section devoted to gaps, some of which will be covered in Chapter 4.

Trading Rules for Gaps

- A *common gap* is small and occurs with low volume and for no specific reason; that is, it is not the result of an obvious, surprising news release. Active traders will take a position counter to the direction of the gap, expecting the move to reverse and fill the gap, at which point they will take profits. If the gap is not filled within a few days, the trade is liquidated.
- A *breakaway gap* is the result of bunched orders at an obvious support or resistance area. When a clear sideways pattern has developed, place a buy order just under the resistance level in order to benefit from the jump in prices (free exposure) when the breakout occurs. If a gap occurs on the breakout, then prices should continue higher.
- A *runaway gap* is often found in the middle of a significant move. It is considered a good point to add to your position because the runaway gap confirms the move and offers additional potential profits.
- An *exhaustion gap* is best traded as it is being filled, and, even at that stage, it is highly risky. Sell during the move upwards, placing a stop above the previous high of the move. If this pattern fails, prices could move higher in an explosive pattern. If you are successive, profits could also be large.

Bulkowski on Gaps

Bulkowski includes an extensive study of breakaway, continuation, and exhaustion gaps. The statistics developed for these three cases all conform to the expected patterns, as shown in Table 3.1. Almost by definition, we expect a breakaway gap, one that occurs when prices move out of a sideways range, to mark the beginning of a new trend, and the exhaustion gap (which actually can't be seen until it reverses) to be the end of a trend. The continuation gap is somewhere in between and is only defined within the context of an existing trend.

The results of the breakaway gap, only 1% and 6% retracements, confirm that the breakouts often continue the trend direction. Strategies that take advantage of this are the *N*-day breakout, swing trading, and pivot point breakouts, providing that the observation period is greater than 40 days, the minimum considered to be a macrotrend.

Spikes

A *spike* is a single, highly volatile day where the price moves much higher or lower than it has in the recent past. A spike can only be recognized one day later because trading range

TABLE 3.1 Percentage of Time Gaps Are Closed within one Week, Based on a Sample of 100 Stocks

Gap Type	Uptrends	Downtrends
Breakaway	1	6
Continuation	11	10
Exhaustion	58	72

of the following day must be much lower. It is easiest to show spikes in markets, such as U.S. 30-year Treasury bonds, that react to frequent economic reports. In Figure 3.14 there is a series of three spikes about four weeks apart.

An upward spike, as shown in Figure 3.14, is always a *local top* because a spike is a day with above-average volatility and must be bracketed by two lower days. In all three cases shown, the spike represented the high price for at least one week. Because the spike is a clear top, when prices begin to rise again, they usually meet resistance at the top of the spike. Chartists draw a horizontal resistance line using the high price of the spike, which encourages selling at that level. After each spike the chart is marked with "failed test," showing the price level where resistance, based on the previous spike, slowed the advance. The spike did not stop the trend, but it did cause a unique pattern.

Quantifying Spikes

A spike has only one dominant feature: a price high or low much higher or lower than recent prices. It must therefore also have high volatility. The easiest way to identify an upside spike is to compare the trading range on the day of the spike to previous ranges and to the subsequent day. This can be programmed in TradeStation by using the *true range* function and satisfying the conditions that the high on the day of the spike is
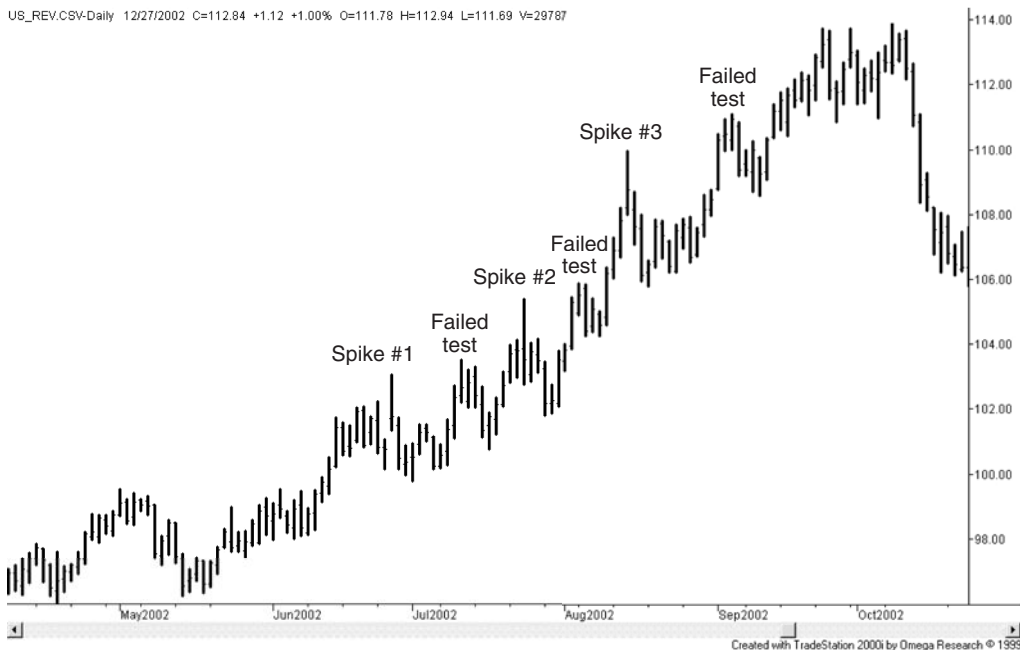
FIGURE 3.14 A series of spikes in bonds. From June through October 2002, U.S. bonds show three spikes that represent local tops. The spikes represent clear resistance levels that cause a unique pattern in the upwards move.

greater than the previous and subsequent highs by the amount of $k \times average\ true\ range$ over n days,

```
Spike = high[1] - highest(high,n)[2] > k*average(truerange,n)[2] and
            high[1] - high > k*average(truerange,n)[2]
```

In this code, *spike* is a logical variable (true-false). A spike that occurs yesterday (where [1] indicates yesterday and [2] two days ago) is tested to see that the high of the spike is greater than the high of the previous n days, greater than the average true range of the same n days by a factor of k, and also greater than the high of today by the same factor k. Note that the use of [2] ends the true range calculation on the day before the spike. The value of k should be greater than 0.75. Spikes satisfying $k > 1$ are more desirable but less frequent.

In Excel, the true range is

$$TRn = Max(Hn - Ln, Hn - Cm, Cm - Ln)$$

where n is the current row, m is the previous row (n-1), and the high, low, and close (H, L, can C) are in columns B, C, and D.

Island Reversals

An *island reversal* or an *island top* is a single price bar, or group of bars, sitting at the top of a price move and isolated by a gap on both sides, before and after the island formation. Combined with high volatility, this formation has the reputation of being a major turning point. The gap on the right side of the island top can be considered an exhaustion gap. In Figure 3.15, showing AMR during the first part of 2003, there is one island reversal in mid-April. This single, volatile day has a low that is higher than both the previous day and the following day. It remains the high for the next week but eventually gives way to another volatile price rise. *Island bottoms* also occur, but are less frequent.

Pivot Point Reversals and Swings

A *pivot point* is a trading day, or price bar, that is higher or lower than the bars that come before and after. If the entire bar is above the previous day and the following day, the pivot point reversal is the same as the island reversal. If it is a very volatile upwards day but the low price is not above the high of the surrounding bars, then it is a spike. If it is not a volatile day, then it is a weaker form of a spike. If you were plotting swing highs and lows, the high of an upwards pivot point reversal day would often become the swing high. It is common to locate a *swing high* by comparing the high of any day with two or more days before and after. The patterns of the days on either side of the high bar are not important as long as the middle bar has the highest high. When more days are used to identify pivot points, these reversals are expected to be more significant; however, they take longer to identify.
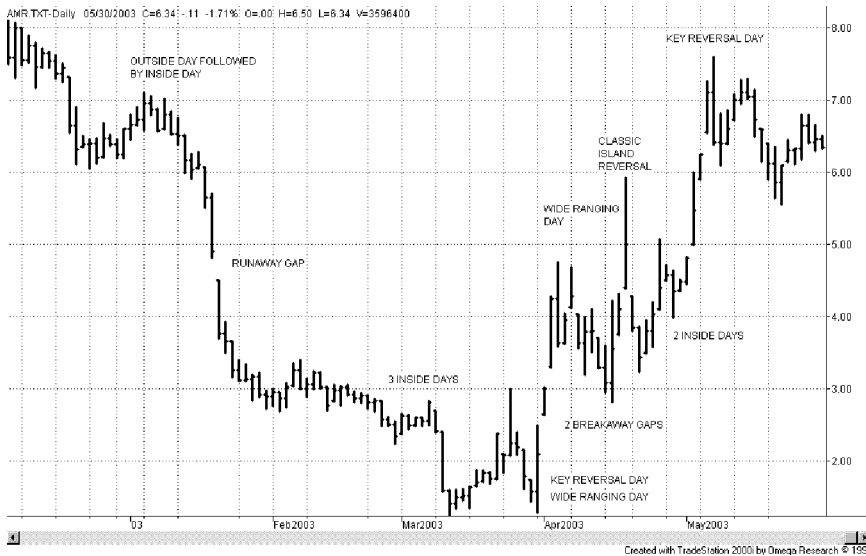
FIGURE 3.15 AMR in early 2003 showing a classic island reversal with examples of other one-day patterns.

According to tests by Colby and Meyers,[13] entries that occur based on a breakout of the highs or lows of the pivot points, called *pivot point channels*, are much more reliable than simply entering in the direction of the reversal based on the close of the last bar of the pivot point formation. For traders not interested in this very short-term strategy, a pivot point may help entry timing for any longer term method.

More recently, Colby[14] tested a *Pivot Point Reverse Trading System*, using the following rules:

- *Buy* (and close out short positions) when a pivot point bottom occurs and the close is higher than the previous close.
- *Sell* (and close out long positions) when a pivot point top occurs and the close is lower than the previous close.

Applying these rules to the Dow Jones Industrials (DJIA) for 101 years ending December 2000 showed nearly 7,000 trades (70 per year) with significant profits. Other tests of pivot points can be found in Chapter 4.

Cups and Caps

Another name given to the pivot point reversals are *cups* and *caps*, each determined by only three price bars, although another formation with the same name, *cup with handle*,

---

[13]Tests of pivot point reversals and pivot point channels can be found in Robert W. Colby and Thomas A. Meyers, *The Encyclopedia of Technical Market Indicators* (Homewood, IL: Dow Jones-Irwin, 1988).

[14]Robert W. Colby, *The Encyclopedia of Technical Market Indicators* (New York: McGraw-Hill, 2003), 510–514.

is similar to a longer-term rounded bottom followed by a sideways or slight downward trend and a breakout to the upside. These two short-term formations are associated with trading rules that are identical to pivot point channels applied to the shortest time frame. Although some literature uses these formations backwards, a cap formation identifies a sell signal when the trend is up, while a cup is a setup for a buy signal in a downtrend. Once an uptrend is clear, a cap formation is found using either the daily closes or daily lows. For any three consecutive days, the middle day must have the highest close or the highest low. In a cup pattern, the middle day must have the lowest low or the lowest close of the 3-day cluster. In both cases, the positioning of the highs and lows of the other two days are not important as long as the middle day is lower for the cup and higher for the cap.

The cup will generate a buy signal if:

- The cup formation is the lowest point of the downtrend
- The buy signal occurs within three days of the cup formation
- The current price closes above the highest high (middle bar) of the cup formation

The signal is false if prices reverse and close below the low of the cup formation, resuming the previous downtrend. This pattern is only expected to forecast a downward price move of two days; however, every change of direction must start somewhere, and this formation could offer an edge. A cap formation is traded with the opposite rules.

Reversal Days and Key Reversal Days

A day in which there is a new high followed by a lower close is a *downwards reversal day*. An *upwards reversal day* is a new low followed by a higher close. A reversal day is a common formation, as seen in Figure 3.16, the Russell 2000 futures. Some of these days are identified; however, you can find many other examples in Figures 3.13 through 3.16. There have been many studies to determine the importance of reversal days for trading, but these are inconclusive. In Chapter 15 there is a detailed study of reversal days, indicating the likelihood of a subsequent price move based on this reversal pattern and other combinations. A reversal day by itself is not significant unless it can be put into context with a larger price pattern, such as a clear trend with sharply increasing volatility, or a reversal that occurs at the highest or lowest price of the past few weeks.

Key Reversals

A *key reversal day* is a more selective pattern, and has been endowed with great forecasting power. It is also called an *outside reversal day*, and is a weaker form of an island reversal. A *bearish key reversal* is formed in one day by first making a new high in an upward trend, reversing to make a low that is lower than the previous low, and then closing below the previous close. It should be associated with higher volatility. Examples of key reversal days can be seen in Figures 3.14 and 3.15. It is considered more reliable when the trend is well-established.
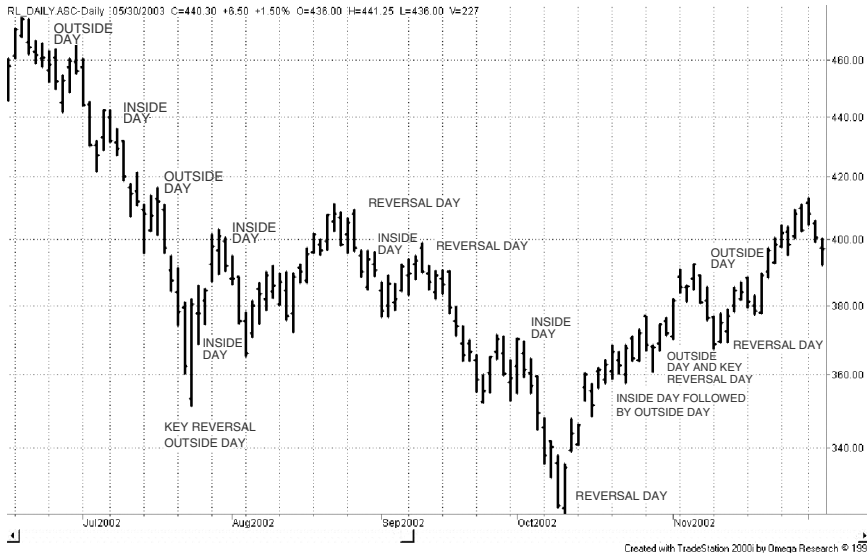
FIGURE 3.16   Russell 2000 during the last half of 2002 showing reversal days, key reversal days, inside days, and outside days.

As with reversal days, studies have shown mixed results using the key reversal as a sole trading indicator. The most complete analysis,[15] similar to others, concluded that the performance was "strikingly unimpressive." Even though tests have not proved its importance, traders still pay close attention to key reversals. Because this pattern has kept its importance, we can conclude that other factors unconsciously enter into the selection of key reversal days for trading. A successful trader's senses should not be underestimated; the extent and speed of the prior trend, a change in liquidity, a quieter market *tone*, or some external news may be essential in confirming the important reversals. The job of a system developer is to find those factors that will turn this pattern into a successful indicator. The best place to start is by assuming the attitude of those traders who see a reversal day as an important pattern. As noted earlier, a comprehensive study of reversal days and other patterns can be found in Chapter 15.

Figure 3.16 shows a number of reversal days during the rapid drop of the Russell 2000 in January 2002. Three patterns of particular interest are the reversal days at the two extreme lows in July and October 2002, and the high in between, during August. Although there are many other reversal days embedded within other parts of the price move, the reversals off the lows are clearly at higher volatility than most other days, and follow very sharp, accelerating price drops. The reversal day that ends the intermediate high during August does not share these attributes; however, it tops a pattern that is not the dominant trend, but an upwards reaction within a previous sustained downtrend. If we focus on the characteristics of those reversal days that mark price extremes, rather than all reversal days, we should expect successful results.

---

[15]Eric Evans, "Why You Can't Rely on 'Key Reversal Days,'" *Futures* (March 1985).

Programming Key Reversal Days

A key reversal day can be recognized and tested using a computer program. In TradeStation's *EasyLanguage* the instructions for downward key reversal are

```
KeyReversalDown = 0;
if close[1] > average(close[1],n) and high >= highest(high[1],n) and
      low < low[1] and close < close[1] then KeyReversalDown = 1;
```

where the first term tests for an uptrend over n-days, the second term tests that the current day is the highest price of the same n-days, the third term verifies that a lower low has occurred, and the last term tests for a lower close. This can be done in Excel by using the *max* and *min* functions instead of *highest* and *lowest*. A TradeStation function to identify key reversals is *TSM Key Reversals* and can be found on the Companion Website along with an Excel spreadsheet of the same name.

Adding a volatility factor, so that the key reversal day has noticeably higher volatility than the previous days seems to select more significant patterns. In the spreadsheet, which uses heating oil from 2005 through 2011 as an example, the basic rules gave marginal gains, but a filter that took only trades where today's true range was greater than $1.5 \times average\ prior\ 20\text{-}day\ true\ range$ was much better.

2-Bar Reversal Patterns

Martin Pring[16] has called attention to a special 2-bar reversal pattern that frequently precedes a strong directional change. This pattern consists of two days that are essentially the mirror image of one another. Consider a market in which prices have been moving steadily higher. The first day of the pattern shows a volatile upwards move with prices opening near the lows and closing near the highs. On the following day, prices open where they had closed, trade slightly higher (nearly matching the previous day's highs), then fall sharply to close near the lows, giving back all of the previous day's move.

Following the 2-bar reversal to the downside, the next few days should not trade above the midpoint of the 2-bar reversal pattern. The smaller the retracement, the more likely there will be a good sell-off.

It is easy to explain the psychology of this pattern. The first bar represents the strong bullish feeling of the buyers, while the second bar is seen as complete discouragement at the inability to follow through to even higher levels. It will take some days before traders are willing to test the highs again. More traders may view this as a potential major reversal. High volume can confirm the reversal. The nature of the move to follow depends on the extent of the previous trend and the volatility. Four key factors in predicting a strong reversal are:

1. Stronger preceding trends
2. Wider, more volatile 2-bar patterns
3. Greater volume than in previous days
4. Smaller retracements following the 2-bar pattern

---

[16]Martin Pring, "Twice as Nice: The Two-Bar Reversal Pattern," *Active Trader* (March 2003).

Wide-Ranging Days, Inside Days, and Outside Days

A *wide-ranging day* is a day of much higher volatility than recent days, but no requirement that it is higher or lower than other days. An *outside day* must have both a higher high and lower low than the previous day. *Inside days* are an example of *volatility compression.* All three patterns are very common but indicate that something special has happened. Examples of these patterns are shown in Figure 3.17, a one-year, active trading period for Tyco ending in July 2000, before any accounting scandal surfaced.

Wide-Ranging Days

A *wide-ranging day* is likely to be the result of a price shock, unexpected news, or a breakout in which many orders trigger one another, causing a large increase in volatility. A wide-ranging day could turn out to be a spike or an island reversal. Because very high volatility cannot be sustained, we can expect that a wide-ranging day will be followed by a reversal, or at least a pause. When a wide-ranging day occurs, the direction of the close (if the close is near the high or low) is a strong indication of the continued direction.

A wide-ranging day is easily seen on a chart because it has at least twice, or three times the volatility of the previous trading days. There is no requirement that it makes a new high or low relative to a recent move, or that it closes higher or lower. It is simply a very volatile day.

Outside Days

An *outside day* often precedes a reversal. An outside day can also be a wide-ranging day if the volatility is high, but when volatility is low and the size of the bar is slightly longer than the previous bar, it is a weak signal. As with so many other chart patterns, if one

FIGURE 3.17 Wide-ranging days, outside days, and inside days for Tyco.

day has an unusually small trading range, followed by an outside day of normal volatility, there is very little information in the pattern. Selection is important.

Inside Days

An *inside day* is one where the high is lower than the previous high and the low is higher than the previous low. That is, an inside day is one where both the highs and lows are inside the previous day's trading range.

An inside day represents consolidation and lower volatility. In turn, lower volatility is most often associated with the end of a price move. After a burst of activity and a surge of upward direction, prices have reached a point where the buyers are already in and the price has moved too far to attract more buyers. Volume drops, volatility drops, and we get an inside day. An inside day is often followed by a change of direction, but that is not guaranteed. We only know that the event that drove prices up is now over. If more news surfaces to ignite prices, the next move could just as easily be up as down

In Figure 3.17 there are two inside days at the price peak on the top left of the Tyco chart. The first inside day is followed by a small move lower, then a small move higher, followed by another inside day. This last inside day precedes a major sell-off. On the right top of the chart there are two inside days immediately before another sharp drop.

Some analysts believe that a breakout from low volatility is more reliable than one following high volatility. For those readers interested in these patterns, a quantitative study of wide-ranging, inside days, and outside days can be found in Chapter 4.

Some Notes about 1-Day Patterns

One-day patterns are very common; therefore, traders tend to be selective about when they are used. Taken as a group, patterns that are repeated frequently are less reliable and need to be combined with other patterns. Those that occur during periods of high or low volatility or volume may also be less dependable.

While a reversal day is clearly a 1-day formation and can be identified at the end of the trading day, and an opening gap is recognized instantly, most other 1-day patterns are not clear until the day after. An upwards spike and a downwards pivot point reversal both require the high of the next day to be much lower than the high of the spike or pivot day; and the island reversal must show a gap on the following day. Although they cannot be used at the end of the day on which they occur, these formations are reasonably timely for an active trader.

CONTINUATION PATTERNS

Continuation patterns occur during a trend and help to explain the stage of development of that trend. A continuation pattern that occurs within a long-term trend is expected to be resolved by continuing in the direction of the trend. If prices fail to move in the direction of the trend following a major continuation pattern, then the trend is considered over. The primary continuation patterns are triangles, flags, pennants, and wedges. The larger formations of these patterns are more important than the smaller ones.

Symmetric, Descending, and Ascending Triangles

Triangles tend to be larger formations that occur throughout a trend. A *symmetric triangle* is most likely to occur at the beginning of a trend when there is greater uncertainty about direction. A symmetric triangle is formed by a price consolidation, where uncertainty of buyers and sellers results in decreasing volatility in such a way that prices narrow to the center of the previous trading range. In Figure 3.18 the symmetric triangle is formed at about the level of the previous support. The breakout from a symmetric triangle often marks the beginning of a longer-term trend.

Formation of a Descending Triangle

Even during a clear downward trend, prices will rally. Because the trend is clear, sellers are anxious to step in and sell these upwards moves, looking for the trend to continue. The top of this mid-trend rally is likely to be the last support point where prices broke out of a previous pattern. In Figure 3.18, the top of the first descending triangle comes very close to the breakout level of the symmetric triangle, and the larger descending triangle towards the lower right of the chart has its high point at the breakout of another descending triangle.

The recent lows of the new trend form a temporary support level, and prices may bounce off that level while short-term traders play for small profits. This action forms a *descending triangle*. As more traders are convinced that prices are still heading lower, rallies off the support level are sold sooner, causing a narrower pattern, until prices

FIGURE 3.18 Symmetric and descending triangles and a developing bear market in gold futures.

finally break below support. The descending triangle is complete. In an upwards trend an *ascending triangle* would be formed.

Size of the Triangles

A triangle should take no less than two weeks to form; however, they can span a much longer period, occasionally up to three months. Larger formations represent periods of greater uncertainty. They may be followed by another symmetric triangle, again indicating that traders are undecided about direction. If the symmetric triangle is resolved in the current trend direction, the trend is in full force, and a large price move is expected.

Triangles can be consistent indicators of investor confidence. Because they reflect human behavior, they are not always perfect in appearance and not always consistent in pattern. It takes experience to identify the formation in a timely manner.

Flags

A *flag* is a smaller pattern than a triangle, generally less than three months for the long-term trader, and is formed by a correction in a bull market or a rally in a bear market. A flag is a congestion area that leans away from the direction of the trend and typically can be isolated by drawing parallel lines across the top and bottom of the formation. At the beginning of a trend, the flags may not lean away from the direction of the new trend as clearly as during a well-established trend. If the first flag after an upwards breakout leans down, it confirms the new upwards trend.

Figure 3.19 shows an assortment of triangles, flags, and pennants. There are two small flags, one in the middle of the chart and one in the lower right, each leaning upwards as expected in a downtrend. A larger flag slightly below center could also have been a symmetric triangle. Both patterns are resolved by a continuation of the trend.

Pennants

Pennants are irregular triangles normally leaning toward the trend, similar to a descending triangle in a downtrend but without a horizontal support line. A typical pennant can be seen in the middle of Figure 3.19. During a sustained trend, triangles are large, clear formations, with horizontal support or resistance lines, while pennants are consolidation formations requiring only that the lines converge. They usually lean in the direction of the trend, but that is not a requirement. A larger pennant should lean in the direction of the trend in a manner similar to a descending triangle; however, a small pennant may serve the same purpose as a flag and lean away from the trend.

Wedges

A pattern that looks as if it is a large pennant, with both sides angling in the same direction, but does not come to a point, is a *wedge*. In an upwards-trending market, the wedge should be rising as shown on the right side of the General Electric chart, Figure 3.20, near

FIGURE 3.19   An assortment of continuation patterns. These patterns are all resolved by prices moving lower. A downward pennant can be found in the middle of the chart.

FIGURE 3.20   Wedge. A weaker wedge formation is followed by a strong rising wedge near the end of 1999 in this chart of General Electric.

the end of 1999. The earlier wedge has nearly a horizontal upper line, bridging the pattern between a wedge and a rising triangle. A rising wedge is formed in the same way as an ascending triangle. Investors, convinced that the share price will rise, will buy smaller and smaller reversals even as prices make new highs. In the end, prices continue in the direction of the trend. In a typical rising wedge the lower line has a steeper angle than the upper line.

The angle of the wedge should be steeper as the trend becomes clear. The earlier wedge formation shown in Figure 3.20 is nearly symmetric. If we study the bigger picture, we can see that the uncertainty at the beginning of the trend is reflected in the symmetric formation, while the rising wedge occurs after the trend is well established and investors anticipate a continuation.

Run Days

Triangles, flags, pennants, and wedges represent the best of the continuation patterns. They can be identified clearly while they are still being formed and the direction of the breakout can be anticipated and traded. Other formations, such as *run days*, are not as timely. A run day occurs when the low of that day is higher than the previous n days, and the high of the day is lower than the subsequent n highs. When it occurs, this pattern confirms that a trend is in effect. The more days used to define the run day, the stronger the pattern. Therefore, a 5-day run day requires 11 days to identify, 5 before the run day and 5 after. Unlike the other continuation patterns, which have a breakout level that can be used as a trading signal, entering a long position after 11 days of a strong upwards move is not likely to be a good entry point. There are no trading rules or trading action associated with run days. They simply confirm what you have already seen on charts—that prices have been trending.

BASIC CONCEPTS IN CHART TRADING

Having covered the fundamental chart patterns, there are some additional concepts that should be discussed in order to keep the proper perspective. Charting involves a great deal of subjective pattern identification; therefore, there may be a choice of patterns within the same time interval. There are also many cases where prices nearly form a pattern, but the shape does not fit perfectly into the classic definition.

Major and Minor Formations

In the study of charting, the same patterns will appear in short- as well as long-term charts. An upwards trendline can be drawn across the bottom of a price move that only began last week, or it can identify a sustained 3-year trend in the financial markets, or a 6-month move in Amazon.com. In general, formations that occur over longer time intervals are more significant. All-time highs and lows, well-defined trading ranges, trendlines

based on weekly charts, and head-and-shoulder formations are carefully watched by traders. Obscure patterns and new formations are not of interest to most chartists, and cannot be resolved consistently unless traders buy and sell at the right points. Charting is most successful when formations are easy to see; therefore, the most obvious buy and sell points are likely to attract a large number of orders.

Market Noise

All markets have a normal level of noise. The stock index markets have the greatest amount of irregular movement due their extensive participation, the high level of anticipation built into the prices, the uncertain way in which economic reports and news will impact prices, and because it is an index. This is contrasted to short-term interest rates, such as Eurodollars, which have large participation but little anticipation because it has strong ties to the underlying cash market, governed by the central bank. The normal level of noise can be seen in the consistency of the daily or weekly trading range on a chart of the Dow or S&P. When volatility declines below the normal level of noise, the market is experiencing short-term inactivity. An increase in volatility back to normal levels of noise should not be confused with a breakout.

   This same situation can be applied to a triangular formation, which has traditionally been interpreted as a consolidation, or a pause, within a trend. This pattern often follows a fast price change and represents a short period of declining volatility. If volatility declines in a consistent fashion, it appears as a triangle; however, if the point of the triangle is smaller than the normal level of market noise, then a breakout from this point is likely to restore price movement to a range typical of noise, resulting in a flag or pennant formation. Both of these latter patterns have uniform height that can include a normal level of noise, but they would not be reliable signals.

ACCUMULATION AND DISTRIBUTION—BOTTOMS AND TOPS

Most of the effort in charting, and the largest payout in trading, goes into the identification of tops and bottoms. For long-term traders, those trying to take advantage of major bull and bear markets, these formations can unfold over fairly long periods. These prolonged phases, which represent the cyclic movement in the economy, are called *accumulation* when prices are low and investors slowly buy into their position, and *distribution* at the top, where the invested positions are sold off.

   The same formations can occur over shorter periods and are very popular among all traders; however, they are not as reliable. There are many top and bottom formations that are popular and easily recognized. In order of increasing complexity, they are the *V*-top or *V*-bottom, the double or triple top or bottom, the common rounded top or bottom, the broadening top or bottom, the head-and-shoulders formation, and the complex top or bottom.

V-Tops and V-Bottoms

The *V-top* (actually an inverted "V"), which may also have a spike on the final day, is the easiest pattern to see afterwards, but the most difficult top formation to anticipate and trade. There have been times, such as in 1974, 1980, and 2000, when the frequency of *V*-tops were deceiving. *V*-tops are preceded by critical shortage and demand, and magnified by constant news coverage. In 1974, it was a combination of domestic crop shortage, severe pressure on the U.S. dollar abroad, and foreign purchases of U.S. grain that combined to draw public attention to a potential shortage in wheat. The news was so well publicized that novice commodity traders withdrew their funds from their declining stock portfolios and bought any commodity available as a hedge against inflation.

It could not continue for long. When the top came in soybeans, silver, and most other commodities, there was no trading for days in locked-limit markets; paper profits dwindled faster than they were made, and the latecomers found their investments unrecoverable. The public often seems to enter at the wrong time.

The most dramatic of all price moves was the technology bubble of the 1990s, ending with a peak in the NASDAQ index during March 2000. As you can see in Figure 3.21, prices rose faster near the end of the bull market, then collapsed just as quickly. It would have been reasonable to expect the move up to end any time after prices penetrated through 3,000, and difficult to expect them to reach 5,000.

Recently, there have been runs in many commodities, but cotton stands out as exceptional. Flooding in both Pakistan and Egypt has greatly reduced the supply causing prices to soar. Figure 3.22 shows prices in February 2011 at levels four times the normal price and a top has not yet formed. Normally, supply shortages in agricultural markets correct

FIGURE 3.21   A V-top in the NASDAQ index, March 2000.

FIGURE 3.22   Cash cotton prices showing V-top in early 2011.

by the next season, but the current run has been so extensive that it may take more time to resolve. Inevitably, it will be solved in the same way, over one or two crop years.

The psychology of the runaway market is fascinating. In some ways, every *V*-top shares a similarity with the examples in Mackay's *Extraordinary Popular Delusions and the Madness of Crowds.* When beef is in short supply, the result of higher feed costs, the consumers do not tend to consider pork, fowl, or fish as an adequate substitute and will accept increased costs longer than expected. This is called *inelastic demand.* As prices near the top, the following changes occur:

- The cost becomes an increasing factor in the standard household budget.
- Rising prices receive more publicity.
- Movements for public beef boycotts begin.
- Grain prices decline due to the new harvest.

This becomes a matter explained by the *Theory of Elasticity.* It can be applied to the 1973 soybean, 1980 silver, and the recent cotton markets. The theory is based on the principle that when prices get high enough, four phenomena occur:

1. Previously higher-priced substitutes become practical (synthetics for cotton, re-claimed silver).
2. Competition becomes more feasible (corn sweetener as a sugar substitute, alternate energy).
3. Inactive operations start up (Southwest gold mines, marginal production of oil).
4. Consumers avoid the products (beef, bacon, silver, cotton).

Consequently, the demand suddenly disappears (the same conclusion arrived at by economists).

Announcements of additional production, more acreage, new products, boycotts, and a cancellation of orders all coming at once cause highly inflated prices to reverse sharply. These factors form a *V*-top that is impossible to anticipate with reasonable risk. There is a natural reluctance to cash in on profits while they are still increasing every day. The situation becomes even more perilous at the end of the move when more investors join the party. These latecomers who entered their most recent positions near the top will show a loss immediately and will need to get out of the trade first; they cannot afford a continued adverse move. Once a reversal day is recognized, there is a mad rush to liquidate. The large number of investors and speculators trying to exit at the same time causes the sharpness in the *V*-top and extends the drop in prices. There is a liquidity void at many points during the decline where there are no buyers and a long line of sellers. A *V*-top or *V*-bottom is always accompanied by high volatility and usually high volume. When the *V*-top is particularly extreme, it is commonly called a *blow-off*. A true *V*-top or *V*-bottom will become an important medium- or long-term high or low for that market.

Two V-Tops in Amazon.com

There is a classic *V*-top in Amazon.com during January 1999, shown in Figure 3.23, and another potential, smaller formation in April. This second one looked as though it was a *V*-top for two days, then quickly disappeared into a broader formation of no particular pattern. A *V*-top cannot be recognized after only a 1-day downturn. The final peak seen in Amazon in late April 1999 is broader than a classic *V*-top but could still be labeled the same.
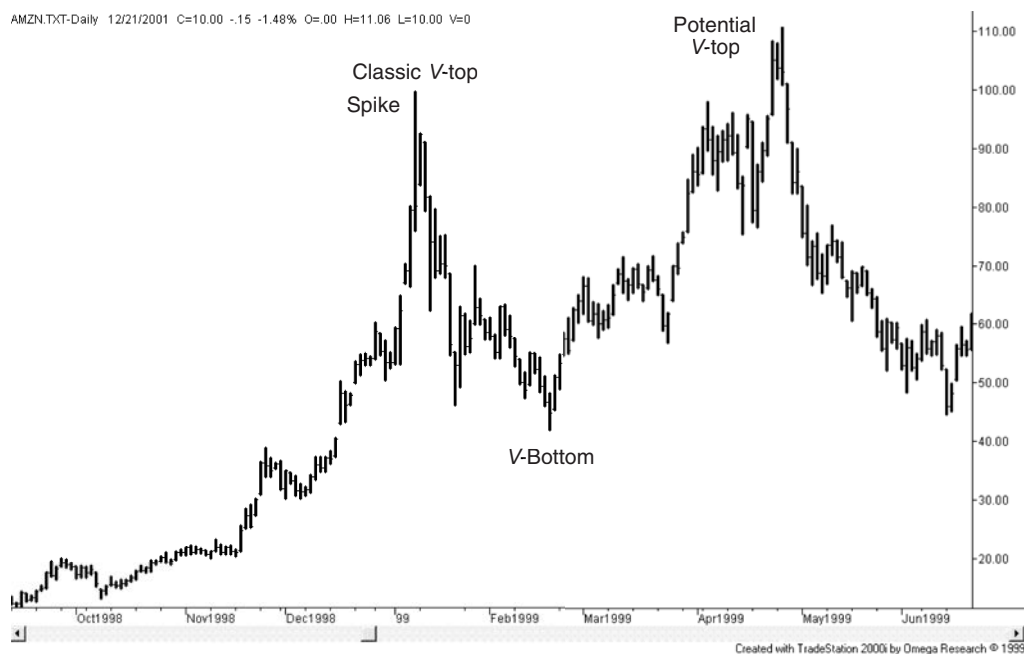
FIGURE 3.23   Classic V-top in Amazon, January 1999, and two other tops.

When trading, you would expect rising prices to fail when they approach the level of a previous clear *V*-top, which forms significant resistance. In Figure 3.23, where prices began the second *V*-top, declined for two days, rallied for the next three days, then dropped sharply for two days, we would normally expect a further decline. In this case, prices made another attempt to break the highs, succeeded, then collapsed. After the last peak, it will be necessary to wait until the price falls below the support level at $75 in order to confirm the downward break, having been fooled on the previous move.

V-Bottoms

V-bottoms are much less common than their upside counterparts. They occur more often in commodity markets where supply and demand can change dramatically and leverage causes surges of buying and selling. Both *V*-tops and *V*-bottoms should be read as a sign that prices have gone too far, too fast. Both buyers and sellers need time to reevaluate the fundamentals to decide where prices should be. *V*-bottoms are usually followed by a rebound and then a period of sideways movement. Two good examples can be found in the crude oil chart, Figure 3.24, and in the stock market crash of October 1987.

Double and Triple Tops and Bottoms

The experienced trader is most successful when prices are testing a major support or resistance level, especially an all-time high in a stock or a contract or seasonal high or low

FIGURE 3.24   Two V-bottoms in crude oil.

in futures. The more often those levels are tested, the clearer they become and the less likely prices will break through to a new level without additional fuel. This fuel comes in the form of higher earnings or a change in the fundamental supply and demand factors.

A *double top* is a price peak followed, a few days or weeks later, by another peak, and stopping very close to the same level. A *double bottom*, more common than a double top, occurs when two price valleys show lows at nearly the same level. Because prices are more likely to settle for a while at a lower price than a high one, prices often test a previous support level causing a double bottom.

Tops and bottoms occur at the same level because traders believe that the same reason that caused prices to fail to go higher the first time will be the reason they fail the second time. The exceptionally high or low prices are the result of speculation rather than fundamentals. In the same way that some stocks will trade at price/earnings ratios far above any rational assessment of business prospects in the near future, commodity prices can be pushed to extremes by crowd psychology without regard to value. Traders, looking for a place to sell an unreasonably high price, target the previous point where prices failed. Although a classic double top is thought to peak at exactly the same price, selling in anticipation of the test of the top may cause the second peak to be lower than the first. Figure 3.25 shows one type of double top in crude oil. While some double tops are two sharp peaks, this one looks as though it was gathering energy. It penetrated slightly above the previous high, but could not sustain higher prices. Double tops are rarely perfect.

FIGURE 3.25 A double top in crude oil.

Double Bottoms

Bottoms are more orderly than tops. They should be quiet rather than volatile. They are caused by prices reaching a level that is low enough for the normal investor to recognize that there is little additional downside potential. Economists might call this the *point of equilibrium.* Neither buyers nor sellers are convinced that prices will continue to move lower. They wait for further news.

Double bottoms will often test the same price level because large-position traders and commercial users of commodities accumulate more physical inventory, or increase their futures position, each time the price falls to their target level. Once prices are low, there is less chance of absolute loss. Selling a double top can be very risky. The greatest risk when buying a double bottom is that your timing is wrong. If prices do not rally soon, you have used your capital poorly.

Cisco shows a double bottom in Figure 3.26, although it lacks the clear decline in volatility that we would like to see, and which accompanies commodities when they reach a price level near the cost of production. The small spikes down show four attempts to go lower, followed by a faster move up. When prices cross above the highs formed between the two bottom patterns, we have a completion, or *confirmation*, of the double bottom.

Traders will start to buy a double bottom when prices slow near previous low levels. They will also look for declining volume or confirmation in the stock price of another related company or a related sector ETF. Waiting for the breakout above the highs of the bottom formation is a safer signal for a conservative trader, but lost opportunity for a more active one.
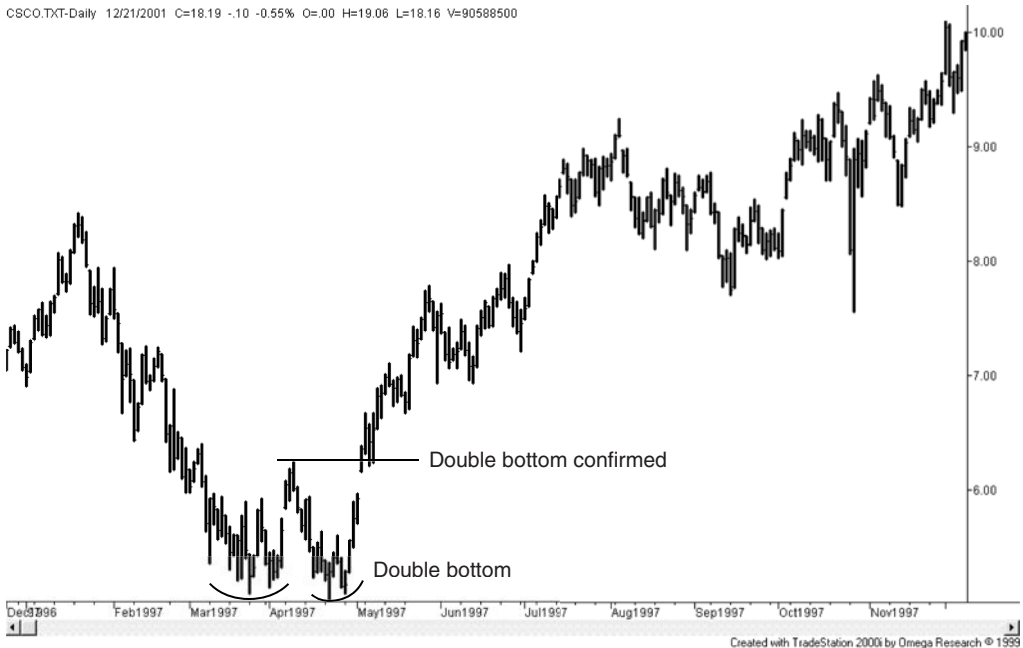
FIGURE 3.26 A double bottom in Cisco.

Triple Tops and Bottoms

Triple tops and *triple bottoms* are considerably less common than double tops and bottoms; however, of the two, bottoms can be found more readily. Figure 3.27 shows a classic triple top in natural gas. A triple top can be formed from a *V*-top, but in this case, the first peak is an island reversal, the second is a spike, and the third an extended top that ends the move.

If we did not have the advantage of seeing the triple top afterwards, each of the individual tops would look as if it were the end of the move. After the first island reversal prices dropped $2; after the second peak there was another large gap down and a 1-day loss of more than $1. High volatility is normally associated with an extreme top. By waiting for a confirmation of a decline after the single or double top, the trade would have been entered $1.50 to $2.50 below the top, and that position would be held while prices reversed to test the highs. Selling tops is risky business.

A triple bottom that can be traded is most likely to occur at low prices and low volatility, much the same as a double bottom. They show an inability to go lower because investors are willing to accumulate a position at a good value. For commodities, it is a good place for a processor to accumulate inventory.

The Danger of Trading Double and Triple Tops

There are many examples of double tops and a smaller number of triple tops. Ideally, there is a lot of money to be made by selling tops at the right place. However, the

NG_REV.CSV-Daily 09/17/2002 C=3.98 +.11 +2.84% O=3.82 H=4.01 L=3.78 V=66195

Created with TradeStation 2000i by Omega Research ® 1999
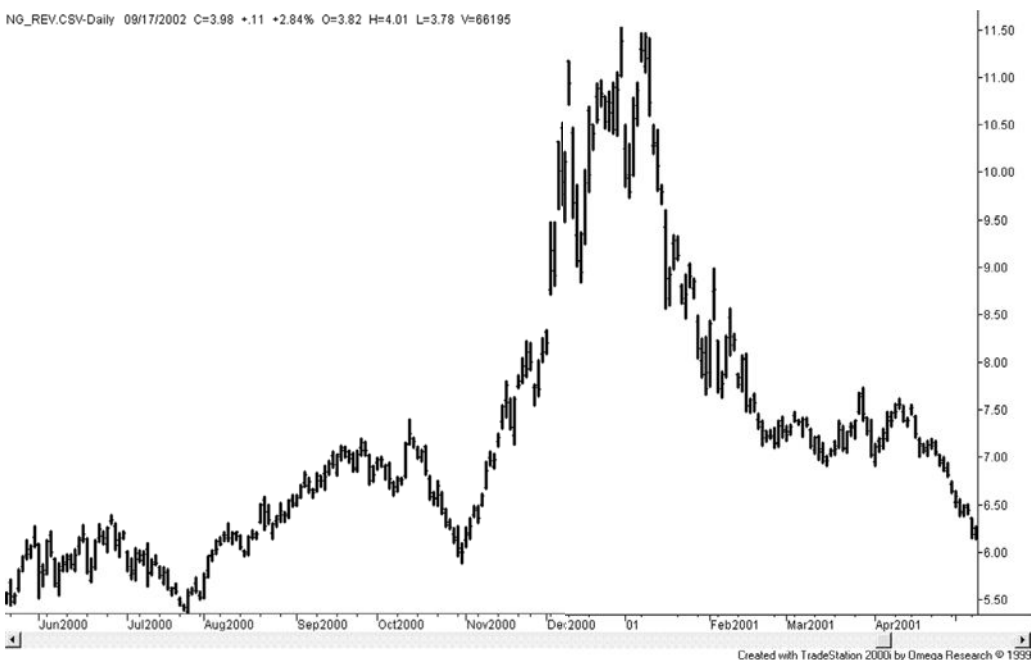
FIGURE 3.27   Natural gas shows a classic triple top.

likelihood of this good fortune happening is less than it appears. Consider why a triple top is so rare. It is because prices continue higher and the potential triple top disappears into a strong bull market pattern.

This happens even more often with double tops. Every time a price pulls back from new highs, then starts moving up again, there is a potential double top. In a prolonged bull market, many double tops disappear in the move higher. Selection of the double top becomes important. This is done using volume, volatility, support, and resistance, and sometimes common sense. These confirming indicators are discussed throughout this book. Until then, it is important to recognize the difficulty of deciding whether the current pattern will be a single, double, or triple top, or simply a pause in a bull market. Although we would all like to be a seller at the highs, these tops are best sold after they are confirmed, that is, after a decline proves that the top has occurred. Even then, a new high should cause a fast exit from the trade.

As with other chart patterns, declining volume would be a welcome confirmation after the formation of the first top and would accompany each additional test of the top.

Extended Rectangle Bottom

Many of the important chart formations can be traded using a penetration of one of the support or resistance lines as a signal. Those with the most potential profitability occur on breakouts from major top or bottom formations. The simplest of all bottom formations, as well as one that offers great opportunities, is the *extended rectangle* at long-term low price levels. Fortunes have been made by applying patience, some available capital, and the following plan:

1. Find a market with a long consolidating base and low volatility. In July 2002 Amazon.com reaches a low with volatility declining. In futures, crude oil remained at low prices for 13 years, as seen in Figure 3.28. The bottom can be confirmed by a decrease in the open interest. When evaluating interest rates, consider using the yield rather than the price.

2. *Buy* whenever there is a test of its major support level, placing a stop-loss to liquidate all positions on a new low price. Increasing volume should confirm the buying, and with futures markets the upside breakout should be accompanied by increasing open interest. For crude oil, the resistance was between $21 and $22 because OPEC had set its OGSP (*Official Government Selling Price*) between $18 and $22.

3. After the initial breakout, buy again when prices pull back to the original resistance line (now a support level). Crude rallied from the lows to just under $40 before pulling back to the old resistance level at $19. Close out all positions if prices penetrate back into the consolidation area and start again at Step 2.

4. Buy whenever there is a major price correction in the bull move. These adjustments, or pullbacks, will become shorter and less frequent as the move develops. They will

Crude Oil Cash Prices
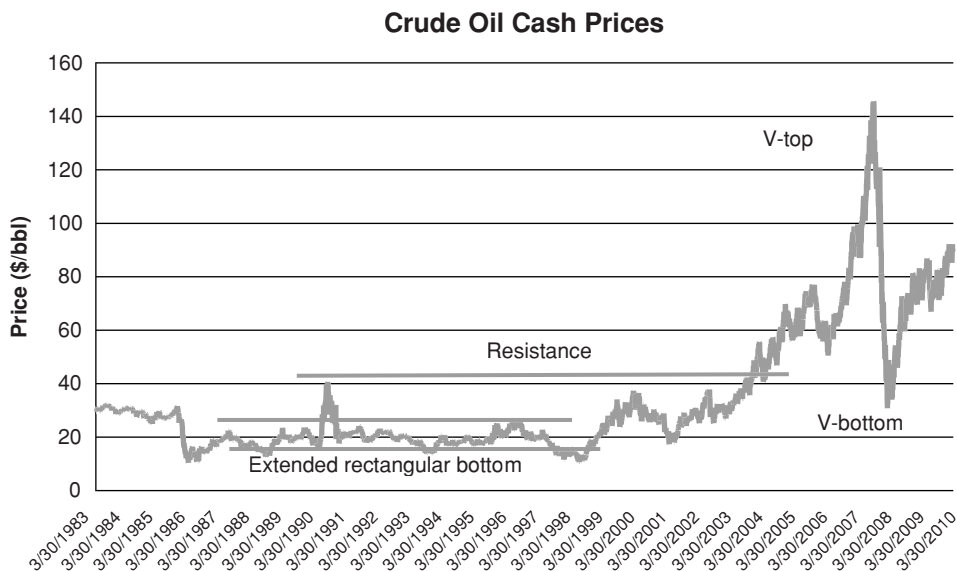
FIGURE 3.28   An extended rectangular bottom in crude oil from 1986 through 1999.

usually be proportional to current volatility or the extent of the price move as meas-
ured from the original breakout.

5. Liquidate all positions at a prior major resistance point, a top formation, or the break-
   ing of a major bullish support line.

Building positions in this way can be done with a relatively small amount of capital
and risk. The closer the price comes to major support, the shorter the distance from the
stop-loss; however, fewer positions can be placed. In his book *The Professional Com-
modity Trader*, Stanley Kroll discussed "The Copper Caper—How We're Going to Make
a Million," using a similar technique for building positions. It can be done, but it requires
patience, planning, and capital. The opportunities continue to be there.

This example of patiently building a large position does not usually apply to bear
markets. Although there is a great deal of money to be made on the short side of the
market, prices move faster and may not permit the accumulation of a large position.
There can also be exceptionally high risk caused by greater volatility. The only pattern
that allows for the accumulation of a large short position is the rounded top, discussed
in the next section. Within consolidation areas for commodities at low levels, there are a
number of factors working in your favor: the underlying demand for a product, the cost
of production, government price support (for agricultural products), and low volatility
itself. There is also a clear support level that may have been tested many times. A careful
position trader will not enter a large short-sale position at an anticipated top when vola-
tility is high, but instead will join the buyers who contribute to the growing volume and
open interest at a well-defined major support level.

Rounded Tops and Bottoms

When prices change direction over a longer time period they can create a rounded top or bottom pattern. A *rounded top* reflects a gradual change in market forces from buyers to sellers. In the stock market it is also called *distribution*. It is a clear sign that any attempt to move prices higher has been abandoned. Rounded tops often lead to faster and faster price drops as more investors liquidate their long positions or initiate shorts.

In Figure 3.29 we see two classic rounded tops in the German DAX stock index. The first is an example of gathering downside momentum as more investors become aware of the decline. Prices drop faster after a break of the double bottom. The rounded top offers a rare opportunity to accumulate a short position with relatively low volatility.

Rounded Bottom

A rounded bottom, similar to a rounded top, is an extended formation where prices gradually turn from down to up. In Figure 3.30 we see a rounded bottom in the Japanese yen followed by a breakaway gap. Similar to the extended rectangle, the rounded bottom offers traders an opportunity to accumulate a large long position. In this case, the sharp rally as prices move through the high of the rounded bottom, followed by a runaway gap, clearly marks the end of the rounded bottom. The breakout can be interpreted as a change in the supply and demand balance. A breakout, whether in stocks or futures, indicates that something new has entered the picture.
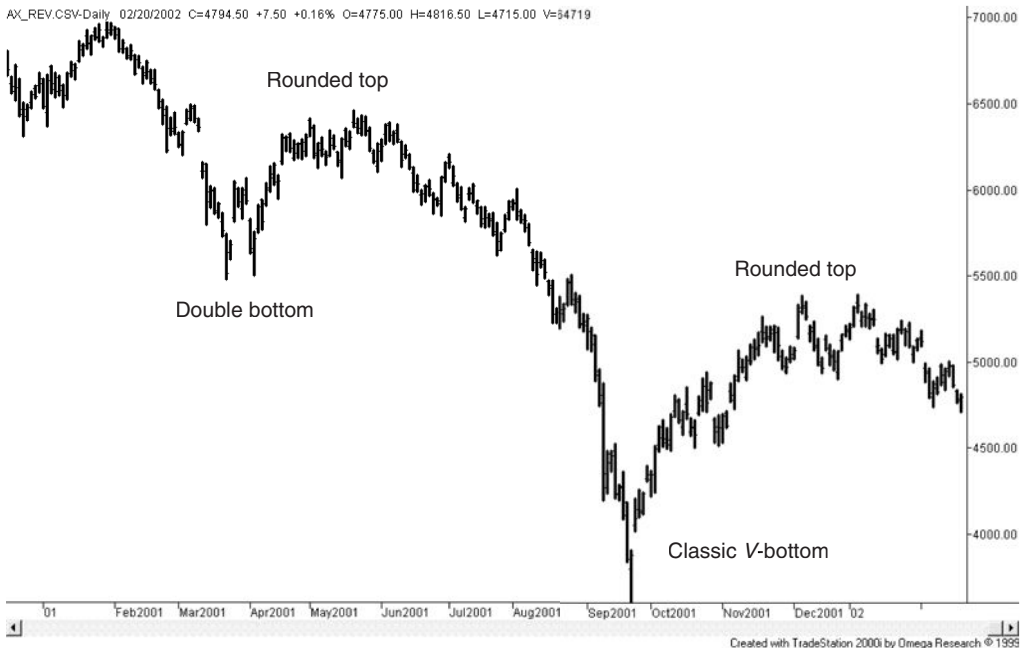
FIGURE 3.29   Two rounded tops in the German DAX stock index.
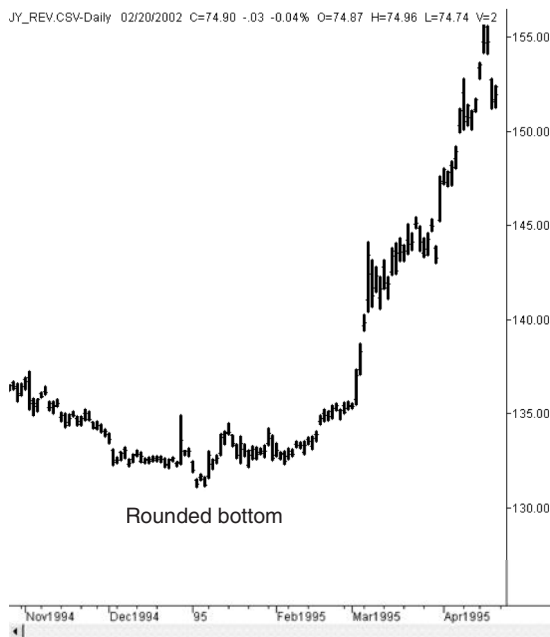
FIGURE 3.30 A classic rounded bottom in the Japanese yen.

Wedge Top and Bottom Patterns

We have seen a wedge formation as a continuation pattern in Figure 3.20, but a large ascending wedge can mark the top of a move and a large descending wedge the bottom. The dominant characteristic of the wedge is that volatility is declining towards the end. In Figure 3.31 there is a declining wedge in the Japanese yen. Volatility compresses until a breakout is inevitable. If the breakout had been to the downside, this wedge would have been interpreted as a continuation pattern. In this example, a breakout in the opposite direction is a strong indicator of a major reversal.

Head-and-Shoulders Formation

The classic top and bottom formation is the *head and shoulders,* accepted as a major reversal indicator. This pattern, well known to chartists, appears as a left shoulder, a head, and a right shoulder, seen in Figure 3.32. The head-and-shoulders top is developed with the following five characteristics:

1. A strong upward breakout reaching new highs on increasing volume. The pattern appears to be the continuation of a long-term bull move.
2. A consolidation area formed with declining volume. This can look much like a descending flag predicting an upwards breakout, or a descending triangle indicating a downwards breakout.

FIGURE 3.31   A large declining wedge
followed by an upside breakout in the Japanese
yen.

FIGURE 3.32   Head-and-shoulders top pattern in the Japanese Nikkei index.

3. Another upwards breakout on continued reduced volume forms the head. This is the key point of the formation. The new high is not confirmed by increased volume, and prices drop quickly.

4. Another descending flag or triangle is formed on further reduced volume, followed by a minor breakout without increased volume. This last move forms the right shoulder and is the third attempt at new highs for the move.

5. The lowest points of the two flags, pennants, or triangles become the *neckline* of the formation. A short sale is indicated when this neckline is broken.

Trading Rules for Head and Shoulders

There are three approaches to trading a head-and-shoulders top formation involving increasing degrees of anticipation:

1. *Wait for a confirmation.*
   a. Sell when the final dip of the right shoulder penetrates the neckline. This represents the completion of the head-and-shoulders formation. Place a stop-loss just above the entry if the trade is to be held only for a fast profit, or place the stop-loss above the right shoulder or above the head in order to liquidate on new strength, allowing a longer holding period.
   b. Sell on the first rally after the neckline is broken. (Although more conservative, the lost opportunities may outweigh the improved entry prices.) Use the same stops as in Step la.

2. *Anticipation of the final shoulder.*
   a. Sell when the right shoulder is being formed. A likely place would be when prices have retraced their way half the distance to the head. A stop-loss can be placed above the top of the head.
   b. Wait until the top of the right shoulder is formed and prices appear to be declining. Sell and place a stop either above the high of the right shoulder or above the high of the head.

   Both steps 2a and 2b allow positions to be taken well in advance of the neckline penetration with logical stop-loss points. Using the high of the head for a protective stop is considered a conservative approach because it allows the integrity of the pattern to be tested before the position is exited.

3. *Early anticipation of the head.*

   Sell when the right part of the head is forming, on the downwards price move, with a stop-loss at about the high of the move. Although this represents a small risk, it has less chance of success. This approach is for traders who prefer to anticipate tops and are willing to suffer frequent small losses to do it. Even if the current prices become the head of the formation, there may be numerous small corrections that will look like the market top to an anxious seller.

Volume was a recognized part of the classic definition of the head-and-shoulders formation and appeared in Robert D. Edwards and John Magee's *Technical Analysis of Stock Trends*, published in 1948. This is no longer considered as important. There are many examples of successful head-and-shoulders formations that do not satisfy the volume criterion. Nevertheless, declining volume on the head or the right shoulder of a top formation must be seen as a strong confirmation of a failing upwards move, and is consistent with the normal interpretation of volume.

EPISODIC PATTERNS

There is little argument that all prices change quickly in response to unexpected news. The transition from one major level to another is termed an *episodic pattern*; when these transitions are violent, they are called *price shocks*. Until the late 1990s, there were very few price shocks in the stock market, the greatest being the one resulting from the terrorist attacks of September 11, 2001. Otherwise, price shocks can be caused by a surprising election result, the unexpected raising of interest rates by the Federal Reserve, the devaluation of a currency by an important Third World nation, sudden crop loss or natural disaster, or an assassination (or what we now call a *geopolitical event*). While price shocks are most common in futures markets, all markets are continually adjusting to new price levels, and all experience occasional surprises. Each news article, government economic release, or earnings report can be considered a mini-shock. A common price shock occurs when a pharmaceutical company's application for a new drug is unexpectedly rejected by the U.S. Department of Agriculture (USDA).

The pattern that results from episodic movement is exactly what one might expect. Following the sharp price movement, there is a period when volatility declines from its highs, narrowing until a normal volatility level is found and remaining at that level. In the Raytheon reaction to 9/11, the upwards price shock, shown in Figure 3.33 is followed by a volatile, unstable few days and then a steady decline in volatility as some level of equilibrium is found. The Raytheon price reacted opposite to most other stocks because it is a defense contractor, and a terrorist attack implies an increased amount of business from the government.

Unless the news that caused the price shock was an error, in which case prices immediately move back to levels prior to the news, prices will settle in a new trading range near the extreme highs or lows. It will take time for the market to absorb the consequences of the news, and many traders will find the risk too high to participate.

Price shocks have become the focus of much analytic work. Because a price shock is an unpredictable event, it cannot be forecast. This has a critical effect on the way in which systems are developed, especially with regard to the testing procedures. We understand at the time of the price shock that the event was entirely unexpected. However, years later, when the same prices are analyzed using a computer program, you might find that a trend or charting pattern *predicted* this move. The analysis records the profits as
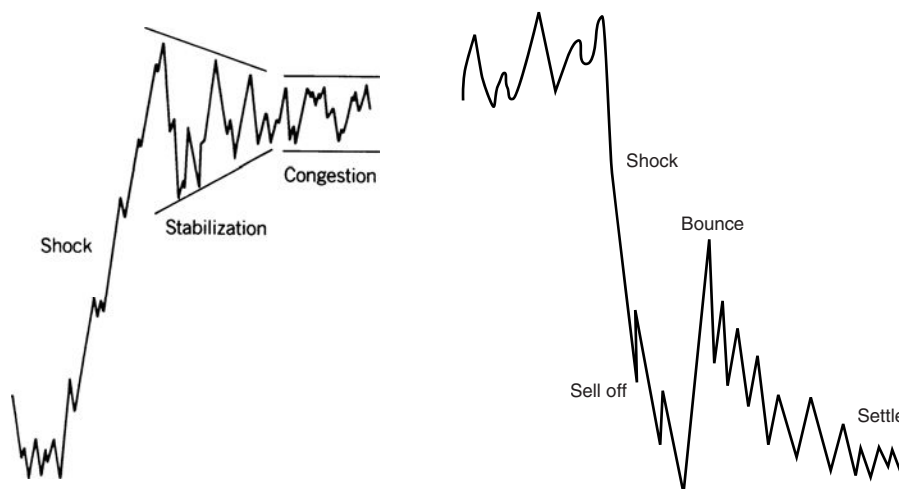
FIGURE 3.33 Episodic pattern shown in an upward price shock in Raytheon following 9/11/2001.

though they were predictable and you are now basing your conclusions on a false premise. These important issues are covered in other parts of this book under the topics "Price Shocks," "Searching for Robustness," and "Optimization," found in Chapters 21 and 22.

PRICE OBJECTIVES FOR BAR CHARTING

Most traders set price objectives and use them to assess the risk and reward of a potential trade. Objectives are most reasonable for short-term trading and successful objectives are based on straightforward concepts and not complex calculations. There is also a noticeable similarity between the price objectives for different chart patterns.

The simplest and most logical price objective is a major support or resistance level established by previous trading. When entering a long position, look at the most well-defined resistance levels above the entry point. These have been discussed in previous sections of this chapter. When those prior levels are tested, there is generally a technical adjustment or a reversal. The more well-established the support or resistance level, the more likely prices will stop. In the case of a strong upwards move, volatility often causes a small penetration before the setback occurs. A penetration of support or resistance, followed by a return to the previous trading range is considered a confirmation of the old range and a false breakout. Placing the price objective for a long position below the identifiable major resistance level will always be safe. The downside objective can be identified in a similar manner: Find the major support level and exit just above it.

When trading with chart patterns, it pays to be flexible. Regardless of which method you use to identify a profit target, be prepared to take profits sooner if the market changes. For example, you have entered a long in IBM at $160 and set your profit objective at

$200. Prices move as predicted and reach $195 when volume starts to drop and the price pattern seems to move sideways. An experienced trader will say "close enough" and take the profit. Profit objectives are not perfect, only good guidelines. If you have set a single price target for a long position, and it falls slightly above a resistance level, then the lower resistance level should be used as the price objective.

One practical solution that will be discussed in Chapter 22 is using multiple profit-targets. Rather than rely on a single point, traders will fan out their target points around the most likely objective, dividing their goal into three or five levels. As each profit-target is reached the risk of the current trade is reduced as is the likelihood of turning a profit into a loss.

While waiting for prices to reach the objective, remember to watch for a violation of the current trend; trend changes take priority over profit objectives. If the trade is successful, and the goal is reached as expected, watch for a new pattern. If prices decline after the trade is closed out, then reverse and break through the previous highs or lows, the position may be reentered on the breakout and a new price objective calculated.

Common Elements of Profit Objectives

Most chart formations have a price objective associated with them. The common ground for all of them is volatility. Each chart pattern is larger or smaller because of the current price volatility; therefore, the price targets derived from these formations are also based on volatility. In general, the *price objective reflects the same volatility as the chart formation* and is measured from the point where prices break out of the pattern.

Profit Targets for Consolidation Areas and Channels

The most basic of all formations is the horizontal consolidation area, bounded on the top and bottom by a horizontal resistance and support lines. There are two possible profit targets, shown in Figure 3.34.

1. For any *horizontal consolidation pattern*, the target is above the breakout of the resistance line at a point equal to the height of the consolidation area (the resistance level minus the support level added to the resistance level). That makes the expected move equal to the extreme volatility of the consolidation area.

2. With *extended rectangular formations*, the upwards profit target is calculated as the width of the consolidation pattern added to the support level.

Although price objective (2) is a well-known and popular calculation, it is unrealistic when the extended formation is very prolonged. The standard calculation, given in (1), is more reasonable. A third objective is more conservative but even more practical:

3. Use the average volatility of the consolidation formation, or reduce the target in (1) above by 20% to remove the extremes from influencing the objectives. A closer price target will be reached more often.
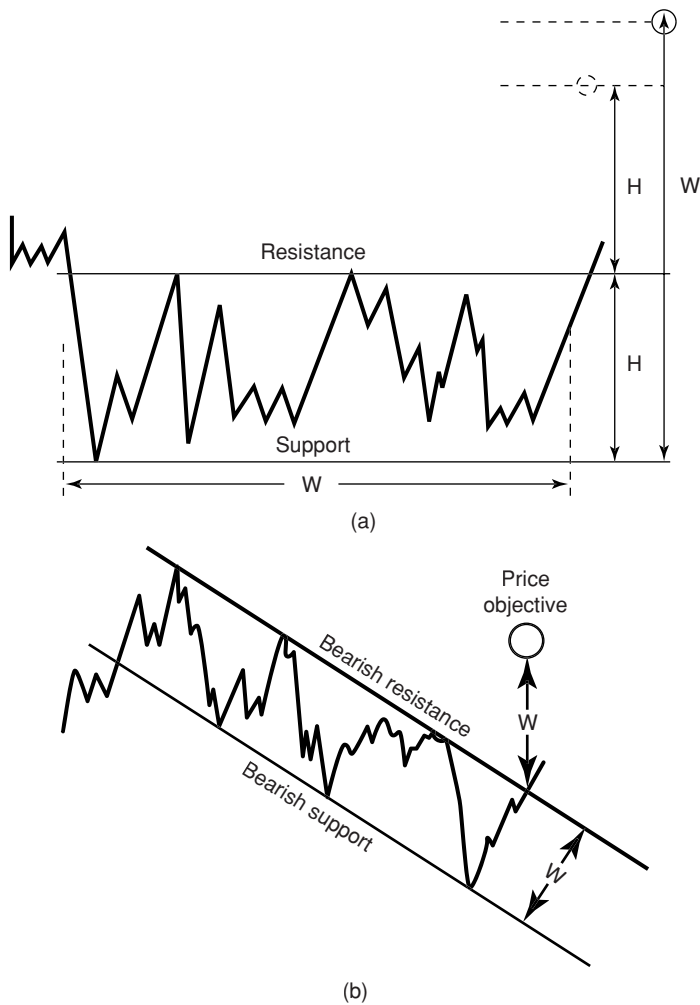
FIGURE 3.34 Price objectives for consolidation patterns and channels. (a) Two objectives for consolidation patterns. (b) Price objective for a channel.

The price objective for a channel is the same as the traditional objective for a horizontal consolidation pattern. Because the channel is at an angle, it is necessary to measure the width of the channel as perpendicular to the angled support and resistance channel lines; then project that width upwards from the point of breakout. The length of the channel does not change the profit target. Again, you may want to make the target slightly smaller than the original channel.

Changing Price Objectives Using Channels

Price objectives can be found as trends change and new channels are formed. Figure 3.35 shows the change from an upwards to a downwards trend. Once a breakout of an upwards channel has occurred (marked "First point of reversal"), we wait until the low is reached
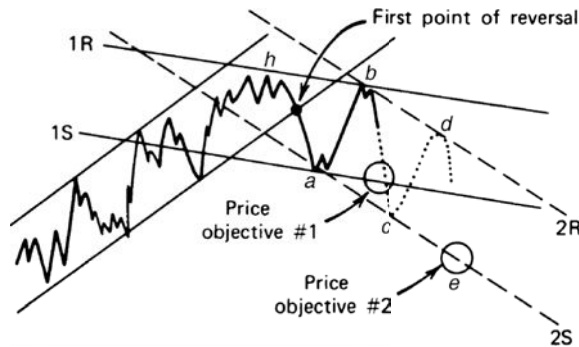
FIGURE 3.35   Forming new channels to determine objectives.

at a, followed by the reaction back up to b. A resistance line, $1R$, can be drawn from the prior high h to the top of the latest move b. A line, $1S$, can be constructed parallel to $1R$ passing through point a, forming the initial downward channel. Price objective 1 is on line $1S$ of the new channel and is used once the top at point b is determined. Price objective 1 cannot be expected to be too precise due to the early development of the channel. If prices continue to point c and then rally to d, a more reasonable channel can now be defined using trendlines $2R$ and $2S$. The support line will again become the point where the new price objective is placed. The upper and lower trendlines can be further refined as the new high and low reactions occur. The primary trendline is always drawn first; then the new price objective becomes a point on the parallel trendline.

Targeting Profits after Tops and Bottoms

Because profit targets are based on the volatility of the underlying pattern, the profit targets for all top and bottom formations will seem very much the same. Looking back at Figure 3.27, natural gas, there is a triple top formation. Between each top is a reversal marking an important support level. The first pullback after the island reversal brought prices to 8.20, followed by a test of the top that formed the second peak. The second retracement stopped at 9.00 and was followed by the third peak. When prices finally drop through the highest support level at 9.00, we can treat it as a breakout and sell short.

If this chart showed a double top, then the point where prices fall below the support between two tops *confirms the top*. Breaking this support level indicates that the topping formation is completed. But this was not a double top; therefore, we can take the lower of the two support levels between the three tops as the major confirmation of the pattern. In the natural gas chart, the lower support was at 8.20. Using either support level gives a measurement of the triple top pattern based on the volatility of prices.

Calculating the Profit Target for a Top Formation
The profit target is found by measuring the height of the top formation and projecting it downwards from the point where the top is confirmed, that is, the break of the support

level. For this example, profit targets will be calculated based on each of the support levels. The highest price of the move is 10.75. Let's examine two profit targets:

1. Using the support level of 9.00, the height of the top is 10.75 – 9.00, or 1.75. Projecting that downwards from the breakout point of 9.00 gives a profit objective of 7.25. The first major pause in the price drop stopped at about 7.00, still showing high volatility.
2. Based on the second support level of 8.20, the height of the top is 2.65, and the profit objective, measured from the break at 8.20, is 5.55. Prices reach 5.55, but only after stalling at about 6.50.

The first target is very achievable and realistic. Prices are very volatile, and a drop of 1.75 could occur very quickly. The second target is less realistic. When targeting a much larger decline, and beginning at a much lower point, it is unrealistic to expect volatility to continue at the same high level. In the decline of natural gas from January through March, volatility also declines, so that by March it appears as though the move is over. Although price targets can often be correct, those that are far away will be less reliable.

Profit Targets after a Bottom Formation
The same principle can be applied to calculate the profit target for bottom formations. The distance from the lowest price of the bottom to the confirmation point is projected upwards from the breakout. This method can be applied to any type of bottom formation. In Figure 3.26, the double bottom in Cisco spanned the price range from about 5.00 to 6.25. The volatility of the bottom pattern, 1.25, is projected upwards from the breakout at 6.25 to get the target of 7.50. Because volatility should expand as prices rise, the exact volatility calculation can be used as a conservative measure.

The Head-and-Shoulders Price Objective
In keeping with other price targets, the head-and-shoulders top has a downside objective, which is also based on its volatility. This objective is measured from the point where the right shoulder penetrates the neckline and is equal to the distance from the top of the head to the neckline (Figure 3.36). For a major top, this goal seems modest, but it will be a good measure of the initial reaction and is generally safe, even if a new high price is reached later.

A very similar example can be found in the Japanese yen (Figure 3.30). The neckline also angles up and to the right, and the price target finds the bottom of the first major support level following the break of the right shoulder. The position of the price objective is so significant that the subsequent drop in prices creates a breakaway gap.

Triangles and Flags
Triangles and *flags* have objectives based on volatility in a manner consistent with other patterns. The triangle objective is equal in size to the initial reaction, which formed the largest end of the triangle (Figure 3.37a). It may also be viewed as a developing channel
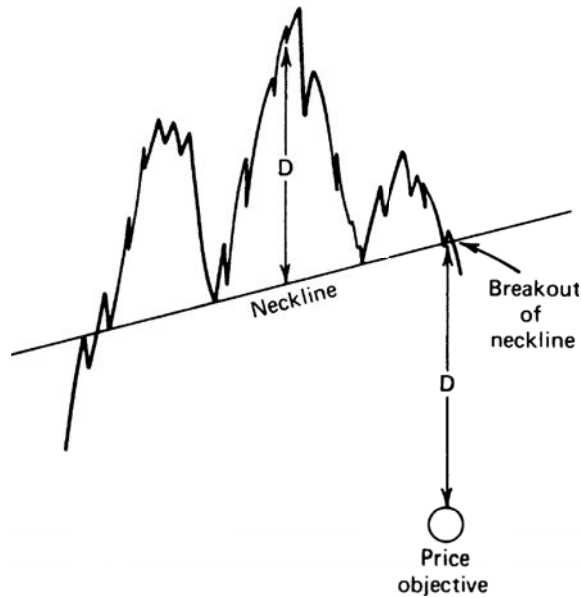
FIGURE 3.36   Head-and-shoulders top price objective.

rather than a triangle, with the ascending leg of the triangle forming the primary bullish trendline. The price objective then becomes the same as those used for channels.

The flag is assumed to occur midway in a price move; therefore, the objective of a new breakout must be equal to the size of the move preceding the flag (Figure 3.37b). Recalling the comments on the problems associated with the decreasing volatility of the triangular formation, the use of the first reaction as a measure of volatility is a safe way to avoid problems. Using this technique with subsequent flags in a bull move will cause objectives to move farther away, becoming unrealistic.

The Rule of Seven

Another measurement of price objectives, the *Rule of Seven*, is credited to Arthur Sklarew.[17] It is based on the volatility of the prior consolidation formation and computes three successive price objectives in proportion to one another. The Rule of Seven is not symmetric for both uptrends and downtrends. Sklarew believes that, after the initial leg of a move, the downtrend reactions are closer together than the reactions in a rising market. Because the downside of a major bear market is limited, it is usually characterized by consolidation. Major bull markets tend to expand as they develop.

---

[17] Arthur Sklarew, *Techniques of a Professional Chart Analyst* (Commodity Research Bureau, 1980).
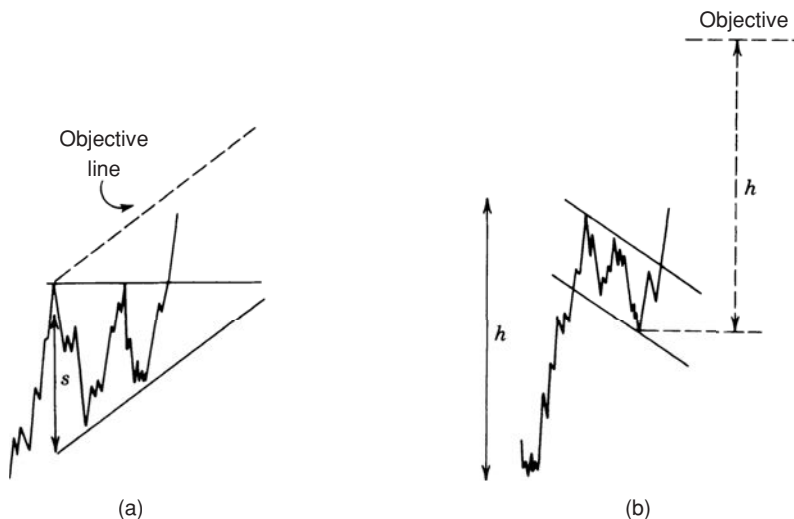
FIGURE 3.37 Triangle and flag objectives. (a) Triangle objective is based on the width of the initial sides. (b) Flag objective is equal to the move prior to the flag formation.

To calculate the objectives using the Rule of Seven, first measure the length L of the initial leg of a price move (from the previous high or low, the most extreme point before the first pullback). The objectives are:

1. In an *uptrend*:

   Upwards objective 1 = prior low + $(L \times 7/4)$
   Upwards objective 2 = prior low + $(L \times 7/3)$
   Upwards objective 3 = prior low + $(L \times 7/2)$

2. In a *downtrend*:

   Downwards objective 1 = prior high − $(L \times 7/5)$
   Downwards objective 2 = prior high − $(L \times 7/4)$
   Downwards objective 3 = prior high − $(L \times 7/3)$

The three objectives apply most clearly to major moves. During minor price swings, it is likely that the first two objectives will be bypassed. In Sklarew's experience, regardless of whether any one objective is missed, the others still remain intact.

IMPLIED STRATEGIES IN CANDLESTICK CHARTS

For a technique that is reported to have been used as early as the mid-1600s, Japanese candle charts were slow to find their way into the western method of analysis. Candle charts can be related to bar charts but offer additional visual interpretation. The *candles*

are created simply by *shading* the piece of the bar between the opening and closing prices: white if the close is higher than the open and black if the close is lower than the open. The shaded area is called the *body* and the extended lines above and below the body are the *shadows*. With this simple change, we get an entirely new way of looking at and interpreting charts. The patterns become much clearer than the Western style of line chart.

Although many candlestick patterns have equivalent bar chart formations, there is an implied strategy in many of them. The following summary uses the traditional candlestick names representing the significance of the formation (see Figure 3.38):
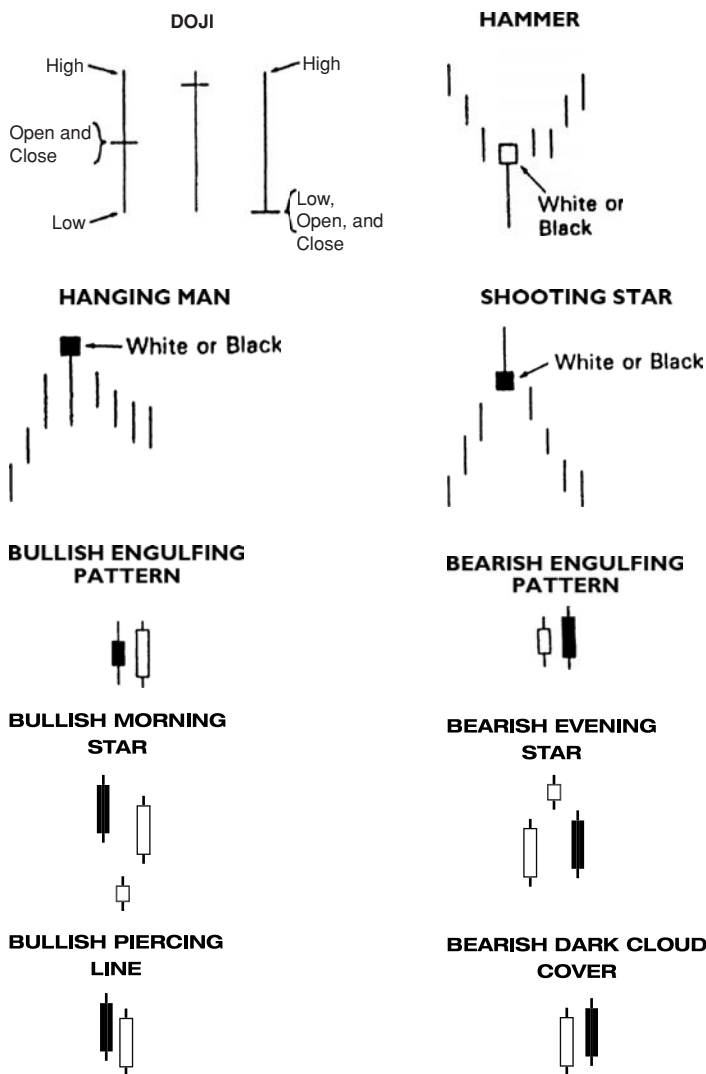
FIGURE 3.38 Popular candle formations.

- *Doji*, in which the opening and closing prices are the same. This represents indecision, a temporary balancing point. It is neither bullish nor bearish. A *double doji*, where two dojis occur successively, implies that a significant breakout will follow.
- *Engulfing patterns* seem at first to be the same as outside days in bar charting, but the pattern only refers to the part of the bar between the opening and closing prices. Engulfing patterns are considered exceptionally strong signals of price change. A *bullish engulfing pattern* has a black candle followed by a white, indicating a wide range with a higher close. The *bearish engulfing pattern* is white followed by black, showing a lower close on the engulfing day.
- *Morning star* and *evening star* are 3-day patterns that show a similarity to an island reversal, but are more specific. In the morning star, a bullish reversal pattern, the first day has a lower close than the open, the second day (called the *star*, similar to the island bottom) has a higher close, and the final reversal day has an even higher close. The bearish reversal is just the opposite, with two higher closes followed by a reversal day with a lower close. If the star is also a doji, then the pattern has more significance.
- *Piercing line* and *dark cloud cover* are bullish and bearish reversals. The piercing line, a bullish reversal, begins with a black candle (a lower close) and is followed by a white candle in which the open is below the previous day's low and the close is above the midpoint of the previous day's body (the open-close range). The dark cloud cover is a bearish formation, the opposite of the piercing line.
- *Hammer*, a bullish reversal signal, showing the bottom of a swing, where the body is at the top of the candle, indicating an upwards change of direction, and the shadow is below the body. The body may be black or white.
- *Hanging man*, a bearish reversal pattern where the body of the candle represents the high of a swing, and the shadow lies below in the direction of the reversal. The body may be black or white.
- *Shooting star*, a bearish signal, also occurs at the top of a swing and has its body at the bottom of the candle with the shadow above. The body may be black or white.

Although these patterns are similar to Western bar chart formations, none of them are exactly the same. The hammer, hanging man, and shooting star are reversal patterns but can only be compared to the simple pivot point where the middle day is higher or lower than the bars on either side. None of these candle formations is exactly the same as a key reversal day or island reversal. The engulfing pattern is stronger than the typical outside day because the spanning of the prior day's range must be done only by the current day's open-close range.

The analysis of candle charts is a skill involving the understanding of many complex and interrelated patterns. For full coverage, Steve Nison's *Japanese Candlestick Charting Techniques*, second edition, is recommended, as well as a selection of newer books, which can be found on Amazon.

Quantifying Candle Formations

The preciseness of the candle formations allow some patterns to be tested. The popular engulfing patterns can be defined exactly for a computer program as

> Bullish engulfing pattern = Previous open > previous close and today's open < previous close and today's close > previous open

> Bearish engulfing pattern = Previous close > previous open and today's open > previous close and today's close < previous open

Another technique uses the shadows as confirmation of direction. We can interpret an increase in the size of the upper shadows as strengthening resistance (prices are closing lower each day); an increase in the size of the lower shadows represents more support. One way to look at this is by defining

Upper shadow (white) = high – close   Lower shadow (white) = open – low

Upper shadow (black) = high – open   Lower shadow (black) = close – low

The sequences of upper and lower shadows can be smoothed separately using a moving average to find out whether they are rising or falling.[18]

A method for determining whether black or white candles dominate recent price movement is to use only the body of the candle, $B = close - open$, and apply a momentum calculation:

$$Body\ momentum = \frac{B_{up}}{B_{up} + B_{down}}$$

where B_{up} = the sum of the days where $B > 0$ (body is white)
          B_{down} = the sum of the days where $B < 0$ (body is black)
                14 = the recommended number of days

When the body momentum is greater than 70, the whites dominate; when the value is below 20 the blacks dominate. These thresholds indicate a built-in upwards bias.

Morning Star and Evening Star

Two formations that are easily programmed are the *morning star* (a bullish signal) and *evening star* (a bearish signal). Using the morning star as an example, the rules call for a long downward (black) candle followed by a lower (the open of the next bar less than the close of the previous long bar), less volatile white candle, and finally an upward thrust shown as a gap up body with the close higher than the open (another white candle).

---

[18] Both "shadow trends" and "body momentum" are adapted from Tushar Chande and Stanley Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994).

When programmed (see *TSM Morning Star* and *TSM Evening Star* in the Companion Website), there are very few signals when we put restriction on the size of the bodies of the three days. Instead, we only required that the body of the first day be greater than the 20-day average, the second day less, and the third day greater. While there are still only a modest number of trades, the S&P performs well on the day following both patterns.

Qstick

As a way of quantifying the Candle formations, Tuschar Chande[19] created *Qstick*, a moving average of the body of the candle. It is intended to be an aid interpreting the charts but has simple trading rules as well.

$$\text{If } Body_t = Close_t - Open_t$$

and

$$Q_t = average(period1, body), \text{ where } period1 \text{ is suggested as 8 days}$$
$$AvgQ_t = average(period2, Q), \text{ where } period2 \text{ is also 8 days}$$

Then the trading rules are

$$\textit{Buy} \text{ when } Q_t \text{ moves above } AvgQ_t$$
$$\textit{Sell} \text{ when } Q_t \text{ moves below } AvgQ_t$$

Pivot Points and Candle Charts

John L. Person suggests that the strategies inherent in candle formations can be combined with support and resistance levels derived from pivot points.[20] He uses the following calculations:

1. Pivot point, $P = (high + low + close)/3$
2. First resistance level, $R1 = (P \times 2) - low$
3. Second resistance level, $R2 = P + high - low$
4. First support level, $S1 = (P + 2) - high$
5. Second support level, $S2 = P - high + low$

Once a key formation for a top or bottom is recognized using candle charts, support and resistance levels calculated based on pivot points can be a strong indication of the extent of the following price move. Person used Dow futures to support his study.

---

[19]Tushar Chande and Stanley Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994).

[20]John L. Person, "Pivot Points and Candles," *Futures* (February 2003).

The Best of the Candles

Bulkowski has summarized his own research in the success of various candles[21] as

- The best-performing candles had closing prices within $\frac{1}{3}$ of the bar low, followed by the middle and high, respectively.
- Candle patterns in a bear market outperform other markets, regardless of the breakout direction.
- Most candles perform best on days with higher volume.
- Candles with unusually long wicks outperform.
- Unusually tall candles outperform.

PRACTICAL USE OF THE BAR CHART

Trends Are Easier to See in Retrospect

As important as it is to identify the direction of price movement, it is much easier to see the trend afterward than at the moment it is needed. There is no doubt that all stocks and futures markets have short-term swings and longer-term bull and bear markets. Unfortunately, at the time you are ready to trade, it is not going to be clear whether the current price trend is a short-term pattern that is about to change or long-term persistent trend experiencing a temporary reversal.

The ease of seeing charts on a screen has made the past patterns clear. It seems natural to expect prices to trend in the future as clearly as they appear on a chart; however, it is not easy to do it in a timely fashion. The eye has a remarkable way of simplifying the chart patterns. The purpose of drawing a trendline is to recognize the direction even though prices can swing violently up and down during that trending interval. A new trend signal to buy or sell always occurs as the trend is changing; therefore, it is at the point of greatest uncertainty.

Success in systematic trading, whether using charts or mathematics, relies on consistency. In the long run, it comes down to probabilities. Success can be achieved by recognizing the trend in 60% of the cases. In a typical trend-following system, because individual profits are much larger than losses, it is only necessary to be correct 30% or 35% of the time.

Long-Term Trends Are More Reliable than Short-Term Trends

Charting is not precise, and the construction of the trendlines, other geometric formations, and their interpretation can be performed with some liberties. When using the simplest trendline analysis, it often happens that there is a small penetration of the channel

---

[21]Thomas Bulkowski, "What You Don't Know About Candlesticks," *Technical Analysis of Stocks & Commodities* (March 2011).

or trendline followed by a movement back into the channel. Some think that this inaccuracy with respect to the rules makes charting useless; however, many experienced analysts interpret this action as confirmation of the trend. The trendline is not redrawn so that the penetration becomes the new high or low of the trend; it is left in its original position.

We must always step back and look for the underlying purpose in each method of analysis, whether interpretive or fully systematic. The trendline is an attempt to identify the direction of prices over some time period. Chartists can use a simple straight line to visualize the direction; they draw the uptrend by connecting the lowest prices in a rising market, even though each point used may represent varying levels of volatility and unique conditions. The chance of these points aligning perfectly, or forecasting the exact price support level, is small. A trendline is simply a guide; it may be too conservative at one time and too aggressive at another; and you won't know until after the trade is completed. Applied rigorously, charting rules should produce many incorrect signals but be profitable in the most important cases. The challenge of the chartist is to interpret the pattern of prices in context with the bigger picture.

Many price moves are called trends, but the most important and sustained trends are those resulting from government policy, in particular those that affect interest rates. Therefore, the most reliable trends are long-term phenomena because government policy develops slowly and is often long-term. During a period of recession, as we saw in 2001 and 2002, the Federal Reserve continued to lower interest rates incrementally, causing a major bull market in all fixed-income maturities. It is easiest to see this trend by looking at a weekly chart of the 10-year Treasury note, rather than an intraday, 1-hour chart. The more detail there is, the more difficult it is to see the long-term trend. Following the subprime collapse of 2008, the Fed and other central banks decided to lower rates to the absolute minimum and keep them there as long as necessary to stimulate the economy. Ultimately, this will result in a protracted bull market in both stocks and commodities.

The average daily impact of the long-term trend on prices is very small. For example, if yields were to drop a staggering 2% in one year, a rise of approximately 16 full points in price, the net effect each day would be a change of .064%, or $\frac{2}{32}$ in price. If prices move nearly one full point, $\frac{2}{32}$ or 1%, each day, that upwards bias would be overwhelmed by the daily market noise. It would be difficult to draw a trendline on a daily price chart until prices had drifted higher for a few months. Using a weekly chart removes much of this noise and makes the trend easier to see.

Multiple Signals

Some of the impreciseness of charting can be offset with confirming signals. A simultaneous breakout of a short-term trendline and a long-term trendline is a much stronger signal than either one occurring at different times. The break of a head-and-shoulders neckline that corresponds to a previous channel support line is likely to receive much attention. Whenever there are multiple signals converging at, or near, a single price, whether based on moving averages, Gann lines, cycles, or phases of the moon, that point

gains significance. In chart analysis, the occurrence of multiple signals at one point can compensate for the quality of the interpretation.

Pattern Failures

The failure to adhere to a pattern is equally as important as the continuation of that pattern. Although a trader might anticipate a reversal as prices near a major support line, a break of that trendline is significant in continuing the downward move. A failure to stop at the support line should result in setting short positions and abandoning plans for higher prices.

A head-and-shoulders formation that breaks the neckline, declines for a day or two, then reverses and moves above the neckline is another pattern failure. *Postpattern activity* must confirm the pattern. Failure to do so means that the market refused to follow through; therefore, it should be traded in the opposite direction. This is not a case of identifying the wrong pattern; instead, price action actively opposed the completion of the pattern. Wyckoff calls this "effort and results," referring to the effort expended by the market to produce a pattern that explains the price direction. If this pattern is not followed by results that confirm the effort, the opposite position is the best option.

Change of Character

Thompson[22] discusses the completion of a pattern or price trend by identifying a *change of character* in the movement. As a trend develops, the reactions, or pullbacks, tend to become smaller. Traders looking to enter the trend wait for reactions to place their orders; as the move becomes more obvious, these reactions get smaller, and the increments of trend movement become larger. When the reaction suddenly is larger, the move is ending; the change in the character of the move signals a prudent exit, even if prices continue erratically in the direction of the trend.

A similar example occurs in the way that prices react to economic reports or government action. The first time the Federal Reserve acts to raise rates after a prolonged decline, the market is not prepared, and interest rate prices react sharply lower. Before the next meeting of the Fed, the market may be more apprehensive, but is likely to be neutral with regard to expectation of policy. However, once there is a pattern of increasing rates following signs of inflation, the market begins to anticipate the action of the Fed. A sharp move in the opposite direction occurs when the government fails to take the expected action.

Bull and Bear Traps

While it is not much of a consolation to those who have gotten caught, a failed downside breakout is called a *bear trap*, and a failed upwards breakout is a *bull trap*. A bear trap occurs when prices fall below a clear support line, generating sell signals. After a few days, prices move back above the support line, often accelerating upwards. A bull trap

---

[22]Jesse H. Thompson, "What Textbooks Never Tell You," *Technical Analysis of Stocks & Commodities* (November/December 1983).

is a failed breakout of a resistance level. In both cases, prices appear to be continuing in the trend direction, but the final picture is a reversal. Although there is no advice on how to avoid bull and bear traps, the failed reversal should be recognized as soon as possible and the position should be reversed. Bull and bear traps often precede significant price reversals.

As with other top and bottom patterns, a confirmation of the bear trap is complete when prices move above the next higher resistance level. In the case of a failed flag formation in a downward trend, prices break lower, as expected, then reverse. The confirmation occurs when prices move above the top of the failed flag pattern. The same principle would be true of other failed chart formations; the failure is confirmed when prices retrace the entire pattern.[23]

Testing Your Skill

Recognizing a pattern is both an art and science. Not everyone has an eye for patterns; others see formations where no one else does. The first decision may be the most important: How much of the chart do you use? It is perfectly normal for different time intervals to show different trends. In some cases, arbitrarily cutting the chart at some previous date might cause a clear trend to disappear. The price scale (the vertical axis) of the chart is another variable not considered by some chartists. When applying methods requiring specific angles, the chart paper is expected to have square boxes. Because of the shape of the box, the formations may appear different from one piece of chart paper, or computer screen, to another.

The timeliness of the pattern identification is the most serious problem. Can the formation be interpreted in time to act on a breakout, or is the pattern only seen afterwards? At different stages of development, the lines may appear to form different patterns. Before using your charting skills to trade, practice simulating the day-to-day development of prices using the following steps:

1. Hold a piece of the paper over the right side of the chart, covering the most recent months, or better still, have someone else give you the partial chart.

2. Analyze the formations.

3. Determine what action will be based on your interpretation. Be specific.

4. Move the paper one day to the right, or have someone else give you the next day's price.

5. Record any orders that would have been filled based on the prior day's analysis. Don't cheat.

6. Determine whether the new day's price would have altered your interpretation and trade.

7. Return to Step 3 until finished.

---

[23]See Christopher Narcouzi, "Winning with Failures," *Technical Analysis of Stocks & Commodities* (November 2001).

This simple exercise might save a lot of money but may be discouraging. With practice you will become better at finding and using formations and will learn to select the ones that work best. Very few traders base their trading decisions entirely on bar charts. Many refer to charts for confirmation of separate technical or fundamental analysis; others use only the most obvious major trendlines, looking for points at which multiple indicators converge. The general acceptance of bar charting analysis makes it a lasting tool.

EVOLUTION IN PRICE PATTERNS

A change has occurred in the stock market because of the S&P 500 index, SPDRs, and other index markets. If you think that stock prices are about to fall because of a pending interest rate announcement by the Fed, you can protect your portfolio by selling an equivalent amount of S&P futures. Afterwards, when you have decided that prices have stabilized, you can lift your hedge and profit from rising prices. It is an easy and inexpensive way to achieve portfolio insurance. You can also speculate in the S&P, NASDAQ, Dow, or sectors, rather than trade individual stocks.

When institutions and traders buy or sell large quantities of S&P futures, the futures price will drift away from the S&P cash index, which represents the weighted average of the actual component stock prices. *Program trading* is the process that keeps the price of futures and ETFs aligned with the cash price of the stocks that comprise those index markets. If you have enough capital, and the difference between the S&P futures price and the S&P cash index is sufficiently large, with the futures higher than the cash, you can sell the S&P futures and buy all of the stocks in the S&P 500 cash index. It is a classic arbitrage that brings prices back together. It is all done electronically in seconds.

But the ability to buy or sell all the stocks in the S&P at the same time has changed the patterns of individual stocks that are part of the S&P index. While at one time these stocks moved largely due to their own fundamentals, they now all move together. It no longer matters that IBM is fundamentally stronger than GE, or that Xerox is at a resistance level and Ford is at support, or even if a company is under investigation. When you buy the S&P futures, you buy all of the stocks at the same time.

Today's technical trader must keep one eye on the individual stock and the other eye on the index. Apple may have moved above its recent resistance level but stopped because the S&P Index is at its own resistance level, and there are more traders watching the S&P than even Apple. In today's market, you can anticipate when a stock will find support and resistance by looking at the S&P chart rather than at the individual stock chart.

Figure 3.39 shows the S&P 500 index, GE, and Exxon (prior to its collapse) over the same period from October 1999 through December 2000. Fundamentally, these three markets have little in common; however, the overall pattern of the three markets is
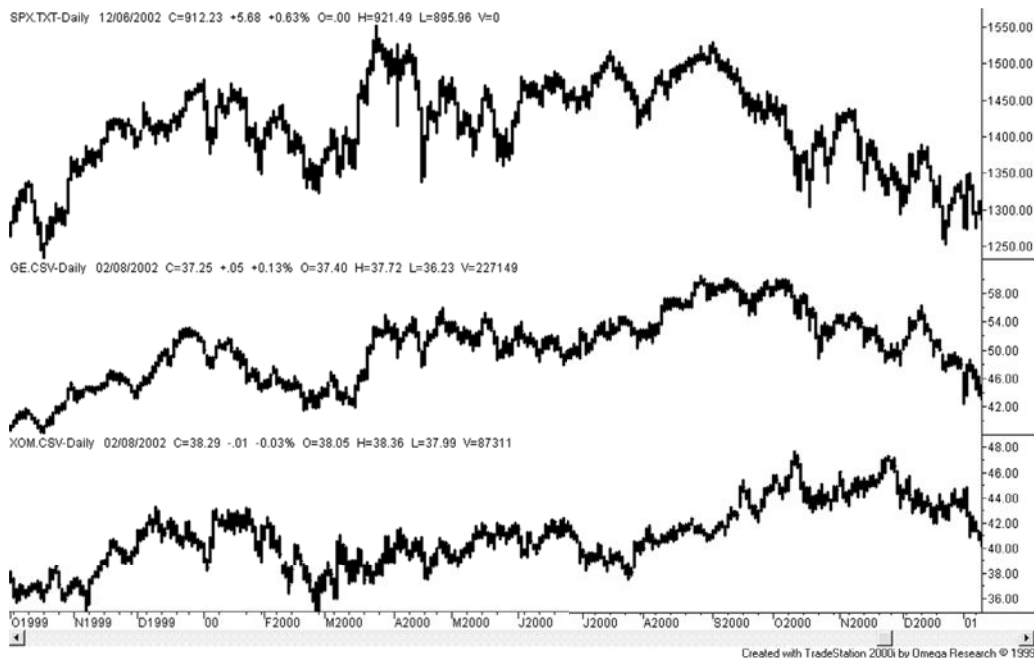
FIGURE 3.39 Similar patterns in the S&P, GE, and Exxon.

remarkably similar, with most tops and bottoms occurring at nearly the same time. Because it is unlikely that the fundamentals of each company would result in such a similar price pattern, we can conclude that the S&P futures, combined with program trading, forces the patterns to be materially the same. This change in the way stocks are traded reduces the ability to get diversification by trading across sectors and increases risk.

Globalization: The Similarity of Asian Markets

There has been a noticeable and justifiable shift to Asian markets during the past five years. Their economies are booming while the United States and Europe are still trying to recover from the financial crisis. Although not all of the Asian stock markets are open to foreign investors, globalization has not passed them by. Figure 3.40 shows the equity index markets for Hong Kong (HSI), Singapore (SSG), Taiwan (STW), the Philippines (PHI), and Malaysia (KLI) as downloaded from Bloomberg. The patterns seem similar but the price levels are very different, making a comparison difficult.

By volatility-adjusting each price series and starting each at 100 on the first date (January 1, 2005), the five series look remarkably the same, as shown in Figure 3.41. It is understandable that, as trading partners, these countries are somewhat dependent upon one another, yet the similarity is surprisingly close. One possible explanation would be the traders. If traders believe that a poor economic sign in one country means that others will also share in bad times, then they sell the equity index markets, or individual
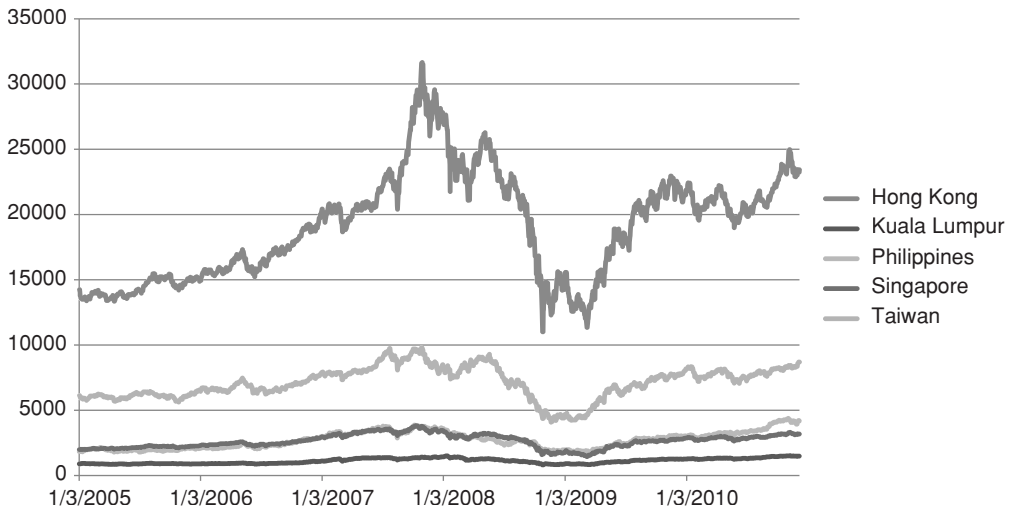
FIGURE 3.40   Equity index prices for five Asian countries.

Data from Bloomberg.

stocks, in each country. That would be similar to Hewlett-Packard announcing worse than expected earnings and having traders sell Dell expecting the same. Often the closer relationships caused by traders show that the movement of money is more important than the fundamentals. This was clearly the case for the subprime collapse in September 2008, when all markets moved the same way as investors withdrew their funds as quickly as possible.
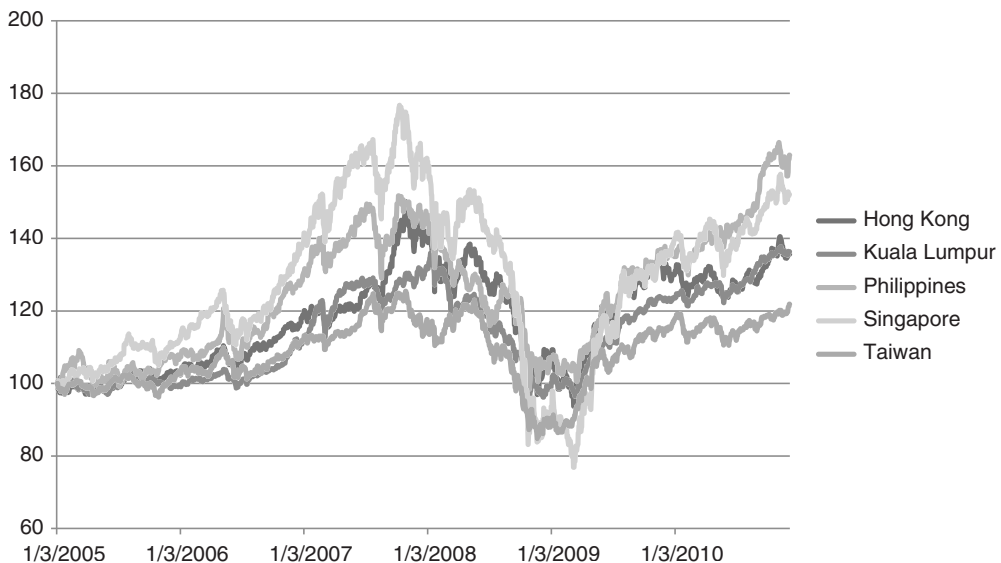
FIGURE 3.41   Asian equity index markets adjusted to the same volatility level and started at the value 100.

Data from Bloomberg.

Charting Systems and Techniques

The continued growth of computer applications has had a great impact on techni-cal trading. The first techniques affected were moving averages and other math-ematical trending methods, then easy-to-program indicators, followed by system-atic optimization. More recently, econometric analysis, cycles, and pattern recognition have been the subject of new development. Many quote services that offer graphics can convert a bar chart to a point-and-figure or candlestick chart at the push of a button. Yet the techniques normally used in classic charting, such as trendlines, channels, and special patterns, are not easily automated because they often depend on the perception of the trader. However, a standard interpretation avoids errors and serves as a useful benchmark.

The systems and techniques included in this chapter are those that might be used by traditional chartists. Many of them are classic methods by famous analysts. They do not all require the use of a chart to be followed, but they are clearly interpretations of natural price patterns. The time that it takes for a price to move from one level to the next is not significant in many of these charting systems; it is only the level itself that is important. The common ground in this chapter is that the methods can be automated. At the end of the chapter is a summary of Bulkowski's work, a study and ranking of most popular chart patterns.

This chapter begins with a review of a few of the earliest attempts at systematic trad-ing. Of course, we have come much further in the 60 years since Dunnigan, and markets have expanded and changed. Yet they are still driven by investors with the same objec-tives. Given our wide range of techniques, tools, and technology, deciding on the most profitable path may be difficult. These first developers struggled with basic concepts and, in many ways, they are the same concepts that we try to resolve now. What appears to be unsophisticated to us now may actually be the key to the best solution. Lest we forget Occam's Razor,

One should not increase, beyond what is necessary, the number of entities re-quired to explain anything.

DUNNIGAN AND THE THRUST METHOD

William Dunnigan's work in the early 1950s is based on chart formations and is purely technical. Although an admirer of others' ability to perform fundamental analysis, his practical approach is contained in the statement:

> *"If the economists are interested in the price of beans, they should, first of all, learn all they can about the price of beans. Then, by supporting their observations with the fundamental elements of supply and demand they will be certain that the bean prices will reflect these things."* [1]

Dunnigan did extensive research before his major publications in 1954. A follower of the Dow Theory, he originally created a *breakaway system* of trading stocks and commodities, but was forced to drop this approach because of long strings of losses. The net results of his system, however, were profitable. He was also disappointed when his "2⅜ Swing Method" failed after its publication in *A Study in Wheat Trading.* But good often comes from failure, and Dunnigan had realized that different measurements should be applied to each market at different price levels. His next system, the *Percentage Wheat Method,* combined a 2½% penetration and a 3-day swing, introducing the time element into his work and perhaps the first notion of *thrust*, a substantial move within a predefined time interval. With the 2½%, 3-day swing, a buy signal was generated if the price of wheat came within 2% of the lows, then reversed and moved up at least an additional 2½% over a period of at least three days.

For Dunnigan, the swing method of charting[2] represented a breakthrough; it allowed each market to develop its natural pattern of moves, more or less volatile than any other market. He had a difficult time trying to find one criterion for his charts that satisfied all markets, or even all grains, but established a $2 swing for stocks where Rhea's *Dow Theory* used only $1 moves. His studies of percentage swings were of no help.

The Thrust Method

Dunnigan's final development of the *Thrust Method* combined both the use of percentage measurements with the interpretation of chart patterns, later modified with some mathematical price objectives. He defines a *downswing* as a decline in which the current day's high and low are both lower than the corresponding high and low of the highest day of the prior *upswing.* If currently in an upswing, a higher high or higher low will continue the same move. The reverse effect of having both a higher high and low would result in a change from a downswing to an upswing. The *top* and *bottom* of a swing

---

[1] William Dunnigan, *Selected Studies in Speculation* (San Francisco: Dunnigan, 1954), 7.

[2] W. D. Gann, *How to Make Profits in Commodities* (Pomeroy, WA: Lambert-Gann, 1976). This book devotes a large section to swing charts and includes many examples of markets prior to Dunnigan's work.

are the highest high of an upswing and the lowest low of a downswing, respectively. It should be noted that an outside or inside day, in which the highs and lows are both greater or both contained within any previous day of the same swing, has no effect on the direction.

In addition to the swings, Dunnigan defines the five key buy patterns:

1. *Test of the bottom*, where prices come within a predetermined percentage of a prior low.
2. *Closing-price reversal*, a new low for the swing followed by a higher close than the prior day.
3. *Narrow range*, where the current day's range is less than half of the largest range for the swing.
4. *Inside range*, where both the high and low fall within the prior range.
5. *Penetration of the top* by any amount, what we now call a *breakout*.

These conditions are reversed for sell patterns. An entry buy signal was generated by combining the patterns indicating a preliminary buy, with a thrust the next day confirming the move. The *thrust* was defined as a price gain that varied with the price level of the market (for 1954 wheat, this was from ½ to 1½¢). Dunnigan's system attempted to enter the market long near a bottom and short near a top, an improvement on the Dow Theory. Because of the risks, the market was asked to give evidence of a change of direction by satisfying two of the first four patterns followed by a thrust on the next day; otherwise, no trade was entered. The same buy and sell signals apply to changes in direction that did not occur at prior tops and bottoms but somewhere within the previous trading range. In the event that all the conditions were not satisfied and prices penetrated either the top or bottom, the fifth pattern satisfied the preliminary signal, and a thrust could occur on any day. This was not restricted to the day following the penetration. If nothing else happened, Dunnigan followed the rules of the Dow Theory to ensure that a major move would not be missed.

Repeat Signals and Double Thrusts

It has been said by followers of Dunnigan's method that his *repeat signals* are the strongest part of his system; even Dunnigan states that they are more reliable, although they limit the size of the profit by not taking full advantage of the trend from its start. Repeat signals use relaxed rules not requiring a new thrust because the trend has already been identified. Two key conditions for repeat buy signals are:

1. A test of the bottom followed by an inside range (interpreted as market indecision).
2. A closing price reversal followed by an inside range.

A *double thrust* occurs when the first thrust is followed immediately by a second thrust; or, after the first thrust, a congestion area develops, followed by a second thrust

in the same direction as the first. Although Dunnigan used a fixed number of points to define his "thrust," today's traders may find today's price move compared to the standard deviation of the daily price changes or the average true range are volatility measures that are more practical for identifying thrusts.

One-Way Formula

Dunnigan worked on what he hoped would be a generalized version of his successful Thrust Method and called it the *One-Way Formula*. Based on his conclusions that the Thrust Method was too sensitive, causing more false signals than he was prepared to accept, he modified the confirmation aspect of the signal and made the thrust into the preliminary signal. He also emphasized longer price trends, which smooth performance and reduce signals.

With the upswing and downswing rules remaining the same, Dunnigan modified the thrust to require its entire range to be outside the range of the prior day. For a preliminary buy, the low of the day must be above the high of the prior day. This is a much stronger condition than his original thrust yet only constitutes a preliminary buy. (It is likely that lower liquidity during the 1950s allowed for more gaps than we have now.) The confirmation requires an additional upthrust after the formation of, or test of, a previous bottom. There must be a double bottom or ascending bottom followed by a thrust to get a buy signal near the lows. If the confirmation does not occur after the first bottom of an adjustment, it may still be valid on subsequent tests of the bottom.

For the One-Way Formula, repeat signals are identical to original confirming signals. Each one occurs on a pullback and test of a previous bottom, or ascending bottom, followed by an upthrust. Both the initial and repeat signals allow the trader to enter after a reaction to the main trend. The Dow approach to penetration is still allowed in the event that all else fails. The refinement of the original thrust method satisfied Dunnigan's problem of getting in too soon.

Updated Trend and One-Way Formula

Ruggiero has interpreted Dunnigan's trend and updated the One-Way Formula[3] so that it can be programmed. An uptrend requires two consecutive days where the highs and lows are both higher, confirmed by prices moving above the high of the current downtrend. As simple as this is, results are good and similar to more complex methods. A program to test this method is *TSM Dunnigan Trend*, available on the Companion Website.

The One-Way Formula is considerably more involved and requires identifying a double bottom, then takes advantage of a short-term uptrend or bounce that follows. A program to test this is *TSM Dunnigan One-Way Formula*, available on the Companion Website.

---

[3]Murray Ruggiero, "Dunnigan's Way," *Futures* (November 1998).

The Square Root Theory

The two previous methods show a conspicuous concentration of entry techniques and an absence of ways to exit. Although it is valid to reverse positions when an opposite entry condition appears, Dunnigan spends a great effort in portfolio management[4] and risk-reward conditions that were linked to exits. By his own definition, his technique would be considered "trap forecasting," taking a quick or calculated profit rather than letting the trend run its course (the latter was called *continuous forecasting*).

A fascinating calculation of risk evaluation and profit objectives is the Square Root Theory. He strongly supported this method, thinking of it as the "golden"[5] key and claiming support of numerous esoteric sources, such as *The Journal of the American Statistical Association*, *The Analyst's Journal*, and *Econometrica*. The theory claims that prices move in a square root relationship. For example, a market trading at 81 (or 9^2) would move to 64 (8^2) or 100 (10^2); either would be one unit up or down based on the square root. The rule also states that a price may move to a level that is a multiple of its square root. A similar concept can be found greatly expanded in the works of Gann (Chapter 14).

NOFRI'S CONGESTION-PHASE SYSTEM

Markets spend the greater part of their time in nontrending motion, moving up and down within a range determined by near-stable equilibrium of supply and demand. Most trend followers complain about the poor performance that results from markets that fail to move continuously in one direction. However, their systems are designed to conserve capital by taking repeated small losses during these periods while waiting to capture the "big move." Eugene Nofri's system, presented by Jeanette Nofri Steinberg,[6] is used during the long period of congestion, returning steady but small profits. Nofri's system does not concern itself with the sustained directional move; therefore the user of the *Congestion-Phase System* can wait to be certain of a well-defined congestion area before beginning a trading sequence.

The basis of the system is a third-day reversal. If prices are within a congestion range and have closed in the same direction for two consecutive days, take the opposite position on the close of day two, anticipating a reversal. If this is correct, take the profits on the close of trading the next (third) day. The concept is that, during a sideways period, sustained runs, either up or down, are unlikely. The Congestion-Phase System is only applied to markets within a trading range specifically defined by Nofri. Users are cautioned not to be too anxious to trade in a newly formed range until adequate time has elapsed or a test of the support and resistance has failed.

---

[4] Each of his writings on systems contained examples of multiple-fund management of varied risk.

[5] Refers to the Greek description of Fibonacci ratios.

[6] A republication of Nofri's method (2010) is available on successincommodities.com.

The *top of the congestion area* is defined as a high, which is immediately followed by two consecutive days of lower closing prices; *the bottom of the congestion* area is a low price followed by two higher days. A new high or low price cancels the congestion area. Any two consecutive days with prices closing almost unchanged (for example, ±2 ticks) are considered as one day for the purposes of the system. In cases where the top or bottom has been formed following a major breakout or price run, a waiting period of 10 additional days is suggested to ensure the continuance of the congestion area and limit the risk during more volatile periods. A congestion area is not formed until both a top and bottom can be identified. Penetration of a previous top and formation of a new top redefine the range without altering the bottom point; the opposite case can occur for new bottoms. If a false breakout occurs lasting two or three days, safety suggests a waiting period of seven days. Logical stops are also possible, the most obvious places being the top and bottom of the current congestion area, but closer stops could be formulated based on price volatility.

Implementing the Congestion-Phase System

When programming the Congestion-Phase System, and most other older strategies, it is likely that not all the rules are as clearly defined as necessary. Some innovation and decisions need to be made. For this method, the greatest uncertainty was defining a "large move" after which we would wait 10 days before looking for new signals. In addition, the "false breakout" needs to be defined clearly and the choice of "two or three" days is taken as "two" days. The "large move" was defined as any net price change (absolute value) over a 10-day interval that was at least two times larger than the average net change over 10 days for all past data. The "false breakout" was any move above or below the congestion levels that reverted back into the congestion zone within two days.

For simplicity, no stop losses were used because trades are held for only one day. However, an interesting characteristic of the pattern of signals should be noted. If we enter a new long position after two days down and the next day is also lower, then we close out the current trade with a loss but also reenter a new long at the same time because the "two day down" rule continues to apply. Once you enter a long (or short) you continue to hold it until you have a 1-day profit or the price moves out of the congestion zone. The TradeStation program, *TSM Nofri Congestion Phase*, can be found on the Companion Website.

Figure 4.1 shows the signals from the program applied to wheat, which was chosen because it would have been a popular market when the strategy was first developed. Note that there were no signals during the rally in late September 2006 due to the "large move" rule. Performance was good for a surprisingly long period, although some markets showed large losses during the 2008 subprime crisis. The concept seems to have a sound basis although it may need to adjust to the extreme volatility in recent years.

The Congestion-Phase System may stand alone as a short-term trading method or can be used to complement any longer technique. After 30 years, it is still unique in the way it defines a sideways range and generates trading signals. There is a shortage of strategies that are not based on trend following, and this method has potential for filling the gap.
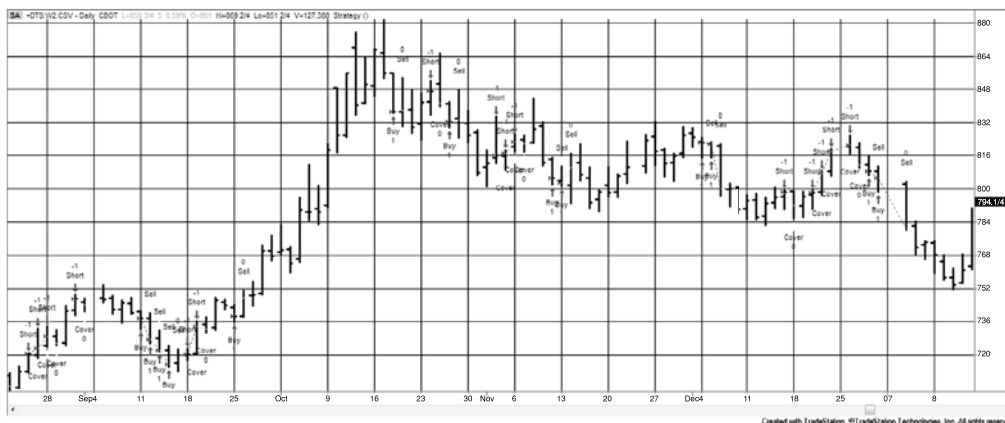
FIGURE 4.1   Nofri's Congestion-Phase System applied to wheat, as programmed on TradeStation.

OUTSIDE DAYS WITH AN OUTSIDE CLOSE

There are numerous chart patterns that can be profitable if they are properly identified and traded consistently. Unfortunately, any one pattern may not appear very often, and traders may become impatient waiting for the opportunities. For others who feel that overall trading success is a combination of small victories, the *outside day with an outside close* is a good place to start.

An *outside day* has the high and low outside the range of the previous day; that is, the high is higher and the low is lower. An outside close is one where the closing price is higher or lower than the prior day's high or low, respectively. This pattern represents a volatile day, usually triggered by news, and is clearly resolved in one direction. If the close was in the direction opposite to a recent price move, it is also a *key reversal day;*[7] however, this method does not attempt to find the current trend. A brief study by Arnold[8] showed that this pattern proved profitable for a small sample of currencies, metals, and financials using the following rules:

1. *Buy* on the close of an outside day if the close is above the prior high; *sell* if the close is below the prior low.
2. *If buying,* place a stop-loss just below the low of the outside day; *if selling,* place the stop just above the high.
3. *Close out* the position on the close three days after entry.

After varying the exits from one to five days, Arnold concluded that this formation predicts reasonably consistent price movements for the next three days.

---

[7]See the discussion of key reversals in Chapter 3.
[8]Curtis Arnold, "Your Computer Can Take You Beyond Charting," *Futures* (May 1984).

Times have changed, and markets are generally noisier and often more volatile. In the 1970s and perhaps into the early 1980s, this pattern was likely to work, but not since the mid-1980s. However, if you reverse the rules and *sell when today's price closes above the previous high* on a volatile day, your results are much better. Arnold held the trade for a fixed number of days, but some conditional exit is more likely to improve results. The program *TSM Outside Day With An Outside Close* is available on the Companion Website. It allows you to test the number of days that the trade is held. Results might be improved by removing trades during periods of low volatility because a wide-ranging day after one with a very narrow range may prove to have no forecasting value.

INSIDE DAYS

Inside days can also be a predictor of direction. Prathap[9] has identified the setup pattern

An upwards day, $C_{t-2} > C_{t-3}$
Followed by an inside day, $H_{t-1} < H_{t-2}$ and $L_{t-1} > L_{t-2}$
Followed by another upwards day, $C_t > C_{t-1}$

as a short-term indicator of a continued upwards move, especially for gold, silver, and crude oil. The reverse is true for downwards moves. A program to test this is *TSM 3-Bar Inside Day*, available on the Companion Website.

PIVOT POINTS

A *pivot point* was defined in the previous chapter as the highest high price or a lowest low in the center of a number of days. Most often there are one to three days on either side of the pivot day. A pivot point can be used in the same way as a swing high or low, except that there is no minimum retracement needed, which adds a greater degree of natural flexibility to the patterns. It is also more restrictive than the swing high, which only needs a high price rather than a reversal after the high occurs; therefore, it introduces a lag as a trade-off for confirmation.

Pivot points can be substituted for swing highs with generally greater success. Their best application is trend following, buying on an upwards move through the previous pivot high and selling on a break through the previous pivot low. Although most uses of pivot points will focus on one to three days on each side of the pivot point, using much longer periods, for example, 10 or 20 days, will give the performance the same appearance as a macrotrend program. A TradeStation program that generates trend signals, *TSM Pivot Point Breakout*, is available on the Companion Website along with an indicator, *TSM Pivot Point*, that plots the pivot points on a price chart, as seen in Figure 4.2.

---

[9]Johnan Prathap, "Three-Bar Inside Bar Pattern," *Technical Analysis of Stocks & Commodities* (March 2011).
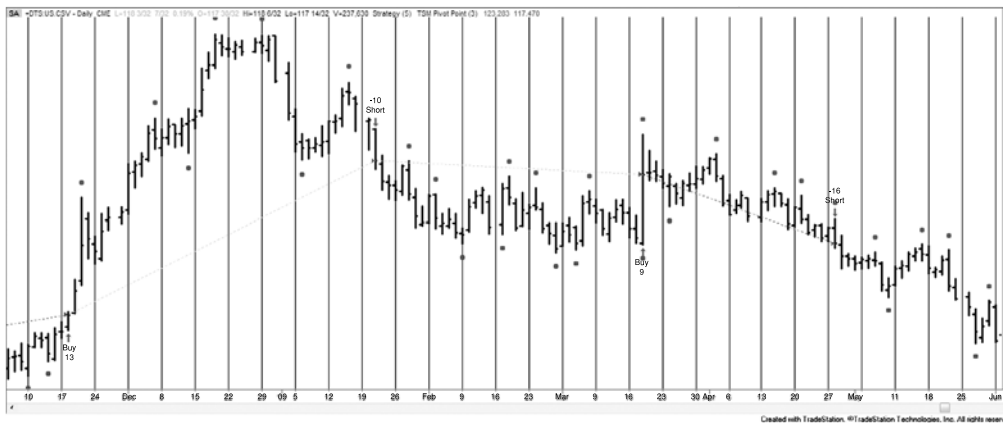
FIGURE 4.2 U.S. 30-year T-bond prices showing pivot points above and below the price and buy and sell signals when there is a penetration of the previous pivot points, based on five days on either side of the pivot point.

ACTION AND REACTION

Fundamentals may be responsible for the ultimate rise and fall of prices, but human behavior creates the patterns that occur as prices find their level of equilibrium. Each move is a series of overreactions and adjustments. Elliott's Wave Principle is the clearest and most well-known of the theories founded entirely on this notion. Frank Tubbs' *Stock Market Correspondence Course* is the first to define the magnitude of these reactions in his Law of Proportion; and, in 1975, the Trident System was based on both the patterns and the size of the action and reaction.

Retracement of a major bull campaign is the most familiar of the market reactions and the one to which almost every theory applies. It is virtually unanimous that a 100% retracement, where prices return to the beginning of the move, encounters the most important support level. The 100% figure itself has been discussed in terms of unity, referring to its behavioral significance. The next most accepted retracement level is 50%, strongly supported by Gann and commonly discussed by experienced speculators. The other significant levels vary according to different theories:

Schabacker accepts an adjustment of ⅓ or ½, considering anything larger to be a trend reversal.

Angas anticipates 25% reactions for intermediate trends.

Dunnigan and *Tubbs* look at the larger ½, ⅔, or ¾ adjustments.

Gann takes inverse powers of 2 as behaviorally significant: ½, ¼, ⅛, . . .

Elliott based his projections on the Fibonacci ratio and its complement (0.618 and 0.382).

Predicting advances to new higher or lower prices is based on prior moves. Gann believed in multiples of the lowest historic price as well as even numbers; prices would find

natural resistance at \$2, \$3, . . . , at intermediate levels of \$2.50, \$3.50, . . . , or at two to three times the base price level. Elliott looked at moves of 1.618 based on a Fibonacci ratio.

Fibonacci Ratios

Along with the most common 1, ½, ⅓, and ¼ retracement values, Fibonacci ratios have the greatest following. Fibonacci ratios are found by dividing one number in the *Fibonacci summation series*

$$1, 1, 2, 3, 5, 8, 13, 34, 55, 89, 144, 233, \dots$$

by the preceding or following value. The series is formed beginning with the values 1, 1 and adding the last two numbers in the series together to get the next value. The numbers in the series, especially those up to the value 21, are often found in nature's symmetry; however, the most important aspect of the Fibonacci sequence is the ratio of one value to the next. Called the *Golden Ratio*, this value F_n/F_{n+1} approaches 1.618 as n gets large. An unusual quality that has drawn attention is that the inverse $F_{n+1}/F_n = 0.618$.

The Golden Ratio has a long history. The great pyramid of Gizah, the Mexican pyramids, many Greek structures, and works of art have been constructed in the proportions of the Golden Ratio. These and other examples are given in Chapter 14 where they are also shown in context with trading systems. At this time it is important to recognize that many analysts who consider human behavior as the primary reason for the size of a price move and their retracements use the Fibonacci ratio 0.618 or, less often, its reciprocal $1 - 0.618 = 0.382$, as very likely targets. Elliott is the most well-known advocate, and applications of his Wave Theory are filled with these ratios. Retracement rules have not been proved scientifically but they are accepted by most traders. In general terms, the retracement theories, or *revelation methods*, can be categorized as either *proportional retracements* or *time-distance goals*. Proportional retracement states that prices will return to a level that is clearly related, by proportion or ratio, to the length of the prior price move. The larger the move, the clearer the retracement. The percentages and ratios expected to be successful are those that are most obvious: 100%, 50%, 33%, and so on, in addition to the Fibonacci ratio 1.618 and its inverse 0.618. The time-distance rule is popularized in the works of Gann (also found in Chapter 14). Gann's retracement objectives can best be thought of as forming an arc of a circle, with the center at the price peak. The goal is satisfied when prices touch any point on the circle.

Practically speaking, it is unrealistic to expect retracement levels to be reached exactly; therefore, when making this fully systematic it is better to allow for the targets to be closer by some percentage in the range of 10% to 15%.

Tubbs' Law of Proportion

The technical part of Frank Tubbs' course in stock market trading is intense chart interpretation. The *Law of Proportion* presented in Lesson 9 of the course is a well-defined
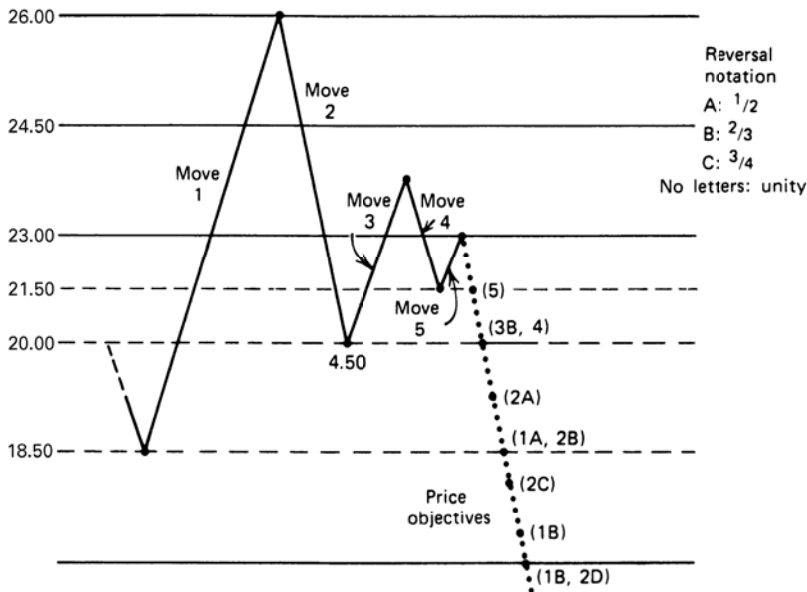
FIGURE 4.3 Tubbs' Law of Proportion.

action-and-reaction law. In cases where the nearby highs or lows of a swing were not broken, Tubbs claims four out of five successful predictions with his principle. The law states:

> *Aggregates and individual stocks tend to run on half, two-thirds, three-fourths of previous moves. First in relation to the next preceding move which was made. Then in relation to the move preceding that.*

Applied to a stock trading at $20, an initial move from $20 to $26 would react ½ to $23, ⅔ to $22, or ¾ to $21.50. Tubbs does allow for traditional price support as a major obstacle to the measured price retracement, and so unity (a 100% retracement) may be added to the three proportions. Figure 4.3 shows subsequent reactions to the stock move just described; the second reversal could be any of three values (or back to major support at $20.00), ending at $21.50, a ¾ reversal. Reversals 3, 4, and 5 are shown with their possible objectives. The last reversal, 5, becomes so small that the major support levels (horizontal broken lines) are considered as having primary significance, along with proportions of moves 1 and 2. Major support at $20.00 coincides with ½ of move 1 and ⅔ of move 2. This would normally be sufficient to nominate that point as the most likely to succeed. Tubbs indicates that these points rarely occur with exactness, but proportions serve as a valuable guideline. The principle is one of reaction in relationship to an obvious preceding action.

Trident

The Trident Commodity Trading System received its fair share of publicity when it was introduced at the beginning of 1975.[10] An article in the 1977 *Dow Jones Commodities Handbook* had an excellent review of the background of the system and some of the conflicts surrounding its presentation and subsequent successes and failures. The system itself is not unique in concept but in its implementation. It is based upon the principle of price action and reaction with formations similar to the waves of R. N. Elliott. For each price move, there is a point of undervalue and overvalue with subsequent reaction, or adjustments, in price as it moves irregularly in the direction determined by the ultimate balance of supply and demand.

The object of the system is to trade in the direction of the main trend but take advantage of the reactions (or waves) to get favorable entry and exit points. The concept of trading with the trend and entering on reactions had been discussed in the context of commodity technical analysis as early as 1942 by W. D. Gann and in the preceding section on action and reaction. As with Gann, the goal is to predict where the reactions will occur and what profit objective to set for each trade. Trident's approach is easy to understand: Each wave in the direction of the main trend will be equal in length to the previous wave in the same direction. The target is calculated by adding this distance to the highest or lowest point of the completed reaction. The determination of the tops and bottoms of the waves is dependent on the time period used; the complex form of primary and intermediate waves, as in Elliott's principle, would hold true with Trident (see Figure 4.4).

Because there are inaccuracies in the measurement of behavioral phenomena, Trident emphasizes the practical side of its theory by offering latitude in its choice of entry and exit points. By entering after 25% of the anticipated move has occurred and exiting 25% before the target, there is ample time to determine that the downward reaction has ended before your long position is taken, and enough caution to exit well before the next reaction. A critical point in each main trend is midway between the start of the move and the target. If the midpoint is not reached, there is a change in direction of the main trend causing a reevaluation of the main trend and the reactions. A change of direction is considered conclusive if a reversal equal in size to 25% of the last reaction occurs during what was expected to be an extension of the main trend. That 25% value becomes the trailing stop-loss on any trade in the event the objective is not reached.

This discussion is not intended to be a complete representation of the Trident System, but a brief description of its essential ideas. The actual system has substantial refinements and subtleties in target selection for major and minor trends and corrective moves; it includes points to reverse positions based on the trailing stop. However, the main premise must hold up if the strategy is to be successful.

In a later bulletin to Trident users, it was suggested that a modification to the system be implemented with respect to money management. Using a technique similar to the Martingale System, each loss is followed by an increase in the size of the position traded.

---

[10]Charles L. Lindsay, *Trident: A Trading Strategy*, reprinted in 1991 by Windsor Books.
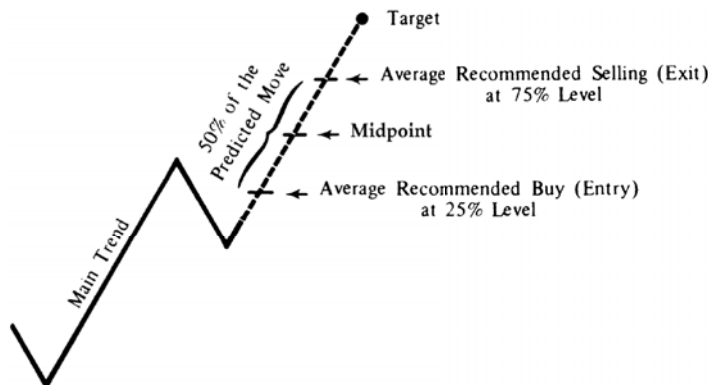
FIGURE 4.4   Trident entry-exit.

The trader only has to continue to extend his positions and stay with the system until he wins. A comprehensive version of this classic gambling approach can be found in the sections "Martingales and Anti-Martingales" and "Theory of Runs," both in Chapter 22. The Trident concepts are all reasonable. They include advance and retracement, trade the trend, don't pick tops and bottoms, take the center out of each move, and use a trailing stop; however, their concept of position size can result in ruin.

An Overview of Percentage Retracements

The last few sections have discussed specific retracement levels advocated by well-known market analysts. This section takes a more general approach to percentage retracements, applying these levels to soybeans and the S&P 500, and draws some conclusions about their use. From previous discussions, the most important reversal that follows a sustained trend is a retracement of 100%, where prices give back all of their gains. A 100% pullback means that the reason for the previous price move has disappeared. This is most common for shorter time periods, where prices are driven by a single news event that turns out to be false, or that the initial reaction to a report was incorrect and prices give back all of their gains, often continuing in the opposite direction.

Confused interpretations of economic reports are quite common. On some days, the government releases multiple reports at the same time, 8:30 Eastern time. These statistics can be new reports of retail sales and revised numbers of past sales. Recently, new housing starts were up 8%, causing a price jump, but then the National Association of Realtors announced a downward revision of 15% for the entire past year. The government revises its previous GDP and other statistics each time it releases new ones. In agriculture, the media reports a major problem with crops each year, the result of too little rain or too much rain. In the end, technology usually wins and there is a record harvest. There are constant rumors of large companies being the target of an SEC probe, or a company correcting an accounting "irregularity." When news turns out to be false, it usually erases the previous gains.

For seasonal commodities, such as crops, there are longer patterns of rising and falling prices. In 1988, a shortfall in the soybean crop with dwindling warehoused stocks caused prices to double. After two years of good harvests, inventories were restored and prices returned to original levels, a 100% retracement. In 2001, the price of soybeans rallied during the growing months, June through August, based on lack of rain. By the end of August, it was clear that the damage was minimal, and prices fell back to the same level as in the spring, a 100% retracement. In 2012 we again have seen one of the greatest bull markets in grain caused by severe drought and anticipation of a harvest that will be at least 10% below normal. How long will that take to retrace? Certainly, news of a good 2013 crop will help, but prices also drop when demand disappears and the high price of corn causes substitution for products such cattle feed, sweeteners, and even reduction in the use of ethanol. The retracement can be faster than we think. Eventually, it will retrace near 100%.

Many patterns used to analyze stock and index prices come from agricultural futures markets. Traders adopt any technique that has worked. The much greater liquidity of soybean futures compared to many individual stocks during the middle of the 1900s makes the futures patterns more reliable. Farmers, grain elevators, and speculators have been doing the same thing since the mid-1800s. Because agricultural products are highly seasonal, and tend to return to the same price levels periodically, they have excellent examples of retracements.

Retracements Less than 100%

There is a significant difference between a *full retracement* (100%) and a *partial retracement.* A full retracement negates the underlying reason for the previous move. But what is the significance of a 50% retracement? Retracements are a common occurrence. They have been compared to the ebb and flow of the tides. Investors buy until they have bought too much, then the sellers come in to correct the overbought situation until the price is back to a level that attracts more buying.

The previous sections have discussed retracements of 50%, 33%, 25%, and 12.5%, as well as 61.8% and 38.2%. The obvious problem is that, if there are so many possible retracement levels, then the price is likely to stop at one of them, even if by chance. Without other information, the most successful retracements are most likely to be the larger ones. Then 100% is the most important, and 50% is the next most likely. After that there is 33% and 25%, each of less importance, with 12.5% too small to consider seriously. Fibonacci ratios are an exception; there seems to be good support for expecting mass behavior to be reflected in these ratios.

Including Fibonacci, the most important retracements are 100%, 50%, and 61.8%. Figure 4.5 shows one of each primary retracement on a weekly soybean chart for a four-year period from 1976 to 1980. Markets that have high volume are most likely to conform to standard retracements. This means that index markets, such as the S&P 500, would also show 50% and 61.8% pullbacks, but individual stocks may not. Mass behavior is a function of broad participation.

S__REV.CSV-Weekly 02/22/2002 C=445.00 +7.75 +1.77% O=439.00 H=447.00 L=435.00 V=106016

Peak2
(exactly twice peak 1)

Peak 1

50% line

38% line

25% line

62% retracement

Starting level

100% retracements

Created with TradeStation 2000i by Omega Research © 1999
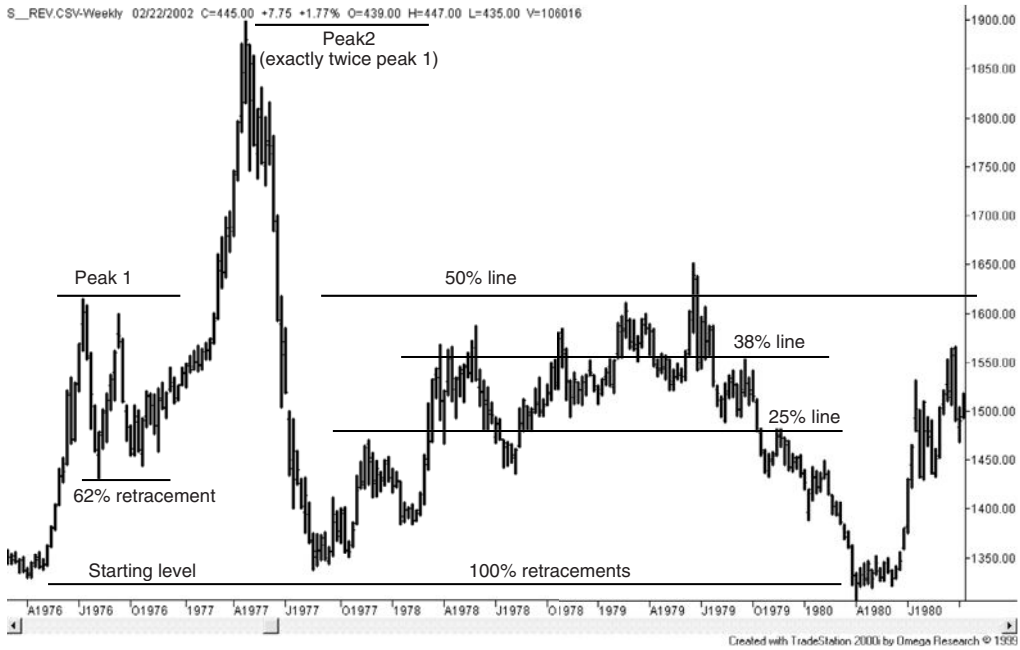
FIGURE 4.5   Soybean retracements in the late 1970s.

S&P Retracement Levels

The S&P has excellent liquidity; therefore, we would expect retracement levels to con-
form to the rule of large numbers. Unlike an agricultural product, or a stock with a strong
seasonal performance, the S&P is not likely to retrace 100% of a longer-term move. We
expect that the core inflation rate, added to the investment bias that exists in the United
States, will cause a steady rise in the overall price of stocks. Figure 4.6 shows the first part
of the bear market that began in 2000. The swing highs and lows are marked with letters
beginning with A and C at the top, with B the low between them. The breakdown of the
support line drawn horizontally from B results in prices reaching D, a decline of 100% of
the range from A to B, followed by a retracement of 50% back to E (support becomes
resistance). Throughout the decline we can find numerous examples of retracement that
conform to the expectations of 100%, 50%, and less important, 62%.

Each retracement level is a trading opportunity. If a rally is expected to stop at a 50%
retracement, a short sale could be triggered automatically at that price. But anticipating
a top and selling into a rising market have a high degree of risk. Price movement is not
so precise that you can anticipate a target with a great degree of confidence. Targeting a
profit level and exiting a trade is considered safe and sensible, because you are removing
risk, but buying when prices are falling quickly is comparable to stepping in front of a
moving train. Entering a new trade on a retracement is considered best when there is a
confirmation that prices have stopped at that retracement level. This may manifest itself

FIGURE 4.6   S&P retracement levels.

as slowing price movement, declining volatility, or declining volume occurring at a point very close to your expected retracement level. If a new low follows on an increase in volume, then you quickly close out the trade and try again later. Common sense is needed in addition to a retracement target.

Clustering

Price movement is a combination of two steps forward and one step backwards. Some serious students of price patterns say that it moves two steps forward and 1.618 steps backwards. Regardless of who is right, prices rarely move in one direction without reversing.

Anticipating the size of a retracement is an attempt to capitalize on the mass behavior of the investors. The continuous flow of funds in and out of the markets reflects the risk tolerance of each participant but also their response to news and the general economy. When seen as the action of a large group, the places where prices stop and reverse, in both their forward movement and backwards steps, seem to cluster at specific levels. Retracements are most important after a sustained price move or during a clear trend.

Interesting observations were made by Tom DeMark[11] about identifying the price move that serves as the basis for measuring retracements. If the market is currently at

[11]Thomas R. DeMark, "Retracing Your Steps," *Futures* (November 1995). Also see Chapter 2 of DeMark, *The New Science of Technical Analysis* (New York: John Wiley & Sons, 1994).

a low, rather than judging the distance of this drop from the most recent swing high, he chooses to look for the highest point that has occurred since the last time the market traded at this low level, thereby eliminating obsolete data. He then finds the most likely retracement points using the Fibonacci ratios 0.618 and 1.618, plus Fibonacci "alternative" ratios 0.382, 0.50, 1.382, 2.236, and 2.618 applied to the difference between the high and low, added to the current low price.

Trading at Even Numbers

It is said that prices advance and decline to even numbers. A stock is more likely to stall at $10 than at $9.25; the price of gold resisted moving below $1000, but once it had traded lower, it struggled to go back above $1000. A study by the New York Federal Reserve confirms the increase in trader activity around even numbers.

It makes sense that investors are more likely to place orders at even numbers. Active traders and longer-term investors do not usually tell a broker to buy at IBM at $153.20, but would more likely buy at $152 or $153. Even more investors would choose $150 or $155. When Martha Stewart placed her now well-known order to sell Imclone stock, it was at $60, not at an odd value.

A good trader can take advantage of this obvious bias for placing orders by avoiding even numbers and looking for *free exposure* when prices move through those levels. Moves through even numbers can be thought of as minor breakouts. If you want to sell a breakout of Imclone short at $60, place your sell order at $60.25 to be ahead of the crowd and take advantage of a fast drop caused by the bunching of orders at even numbers.

CHANNEL BREAKOUT

The classic upwards *channel* is formed by drawing a straight line along the bottom points of an upwards trend, then constructing a parallel line that touches the extreme high price of that same time interval, forming an envelope, or channel, around a price move. For a downward channel, the trendline is first drawn through the high points of the declining price pattern; then a parallel line is drawn through the lowest low price of that interval. It is easy to do this with a chart and a ruler, even easier with drawing tools on a screen, but not as simple to transfer this concept to a computer program. Because a channel breakout is basic trading strategy, an automated version may prove useful for identifying key market turning points. Two different approaches can be used:

1. *Swing high and low points*
   a. Locate the swing high and low points using the program discussed at the beginning of Chapter 5. Alternately, pivot points can be used.
   b. The most recent trendline connects the two most recent high swing points or low swing points.
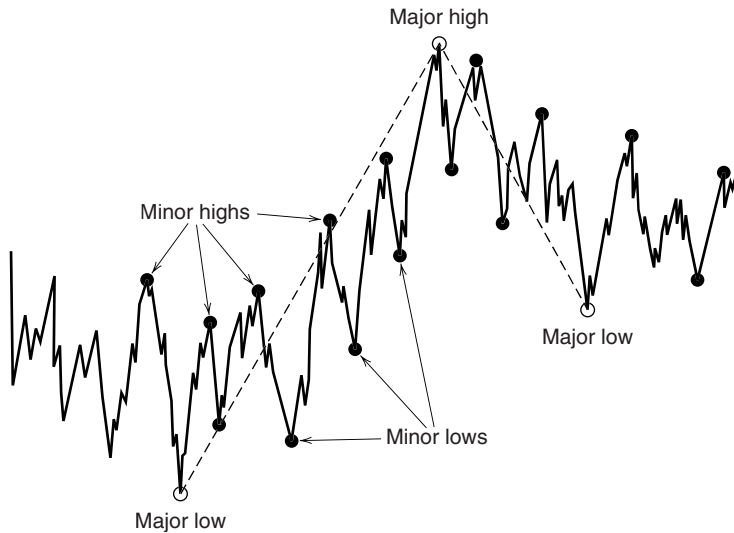
FIGURE 4.7   Major and minor swings.

    c. For a longer uptrend, use a least-squares regression of the most recent 3, 4, 5, etc. low swing points. Choose the number of points that give the best fit, provided the slope is angling upwards.

    d. Draw a parallel line across the highest high of the same time interval using the same slope as the regression line.

    e. Using two different swing percentage criteria, faster and slower trendlines can be found (see Figure 4.7).

2. *Use a straight-line fit*

    a. Select the data beginning with the last swing high or swing low, based on observation or the swing program.

    b. Use the closing prices for Y (the dependent variable) and sequential integers 1, 2, 3, . . ., t for the values of X (the independent variable) through today. It would be easier if we could use the date instead of a sequential number, but the gaps due to weekends would cause an incorrect answer. Solve for a and b, the slope and y-intercept, using the regression data analysis function in Excel located in the menu at Data/Data Analysis/Regression. Once you have the a and b values, any point on the line can be found using

$$\text{Straight line values, } Y = a \times X + b$$

    c. Create trendlines parallel to the regression line by finding the distance BL from the line to the lowest point below the line, and BU, from the line to the highest point above the line. Then the upper trendline (resistance) and the lower line (support) are

$$Y_U \, (upper) = a \times X + b + BU$$
$$Y_L \, (lower) = a \times X + b - BL$$

d. *Project the bands one period ahead.* In order to know whether the next price has broken through the channel, indicating a change of trend, we project the channel one period ahead using the slope value, a:

Projected upper channel band at $n + 1 = a \times X + b + BU + a$
Projected lower channel band at $n + 1 = a \times X + b - BU + a$

If the trend is up (the slope $a > 0$) and the next price, either the close or the low, is *below* the projected *lower* band, then the trend has turned from up to down. If the trend is down ($a < 0$) and the next price, either the close or high, is greater than the projected upper band, then the trend has turned up. When the slope, a, is very near zero, we have a sideways channel, but the same rules still apply. The use of the closing price to decide the breakout, rather than the high or low, is more conservative.

Because a major channel is considered a strong chart formation, prices that approach the channel, but have not penetrated the band, would be candidates for a countertrend entry. For example, if the trend is down and prices come within 15% of the upper band (based on the channel width), we would enter a new short position (see Figure 4.8) or
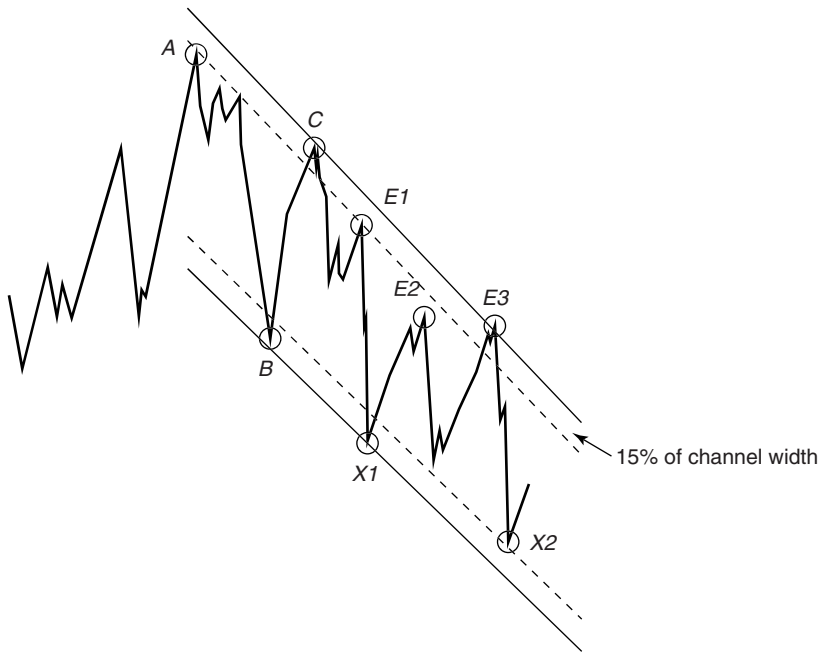
FIGURE 4.8 Entering near the top of a declining channel.

add to existing short sales. We do not necessarily want to cover those existing shorts at the bottom of the channel, especially if the downtrend is severe; however, this technique offers a clear and safe way to scale into a trade with more than one entry point. The trade is closed out if the price breaks above the upper channel line in a downtrend or the lower channel line in an uptrend. If the trend is sideways (the slope is near zero) then exiting shorts and reversing to a long position is the preferred strategy. One note of caution: all trends turn sideways as they reverse direction. For a sideways market the rate of change of the price must also be small.

A minor channel can also be used to signal buys and sells within a major channel. By finding swing points using a smaller minimum swing value, we can wait for a break of the minor channel to signal a new trend direction within the major channel. This gives an added confirmation that prices are moving in your direction before entering.

MOVING CHANNELS

Channels are frequently constructed as moving bands around prices. Some of these, such as those using a standard deviation, can claim statistical significance (these are discussed in Chapter 8, in particular the section "Bollinger Bands"). A simple mathematical way of representing a moving channel might use the average of the high, low, and close (M) to designate the center of the channel (a substitute for the straight line); the upper and lower bands are constructed using the average daily range (or true range), R. The midpoint of the price move M and range R can be calculated over the past n days as

$$M_t = \frac{1}{3n} \sum_{i=t-n+1}^{t} (H_i + L_i + C_i)$$

$$R_t = \frac{1}{n} \times \sum_{i=t-n+1}^{t} (H_i - L_i)$$

Then the upper and lower channel bands add and subtract ½ R to the midpoint M, and the forecast for the next day will project the path of the midpoint and apply a multiple of the range (f) for scaling,

$$U_{t+1} = M_t + (M_t - M_{t-1}) + f \times \frac{R_t}{2}$$

$$L_{t+1} = M_t + (M_t - M_{t-1}) - f \times \frac{R_t}{2}$$

A long position is taken when the new price $p_{t+1} > U_{t+1}$; a short is taken when $p_{t+1} < L_{t+1}$. If a *channel profit objective* is needed, it can be calculated at a point equal in distance to the channel width from the channel breakout as follows
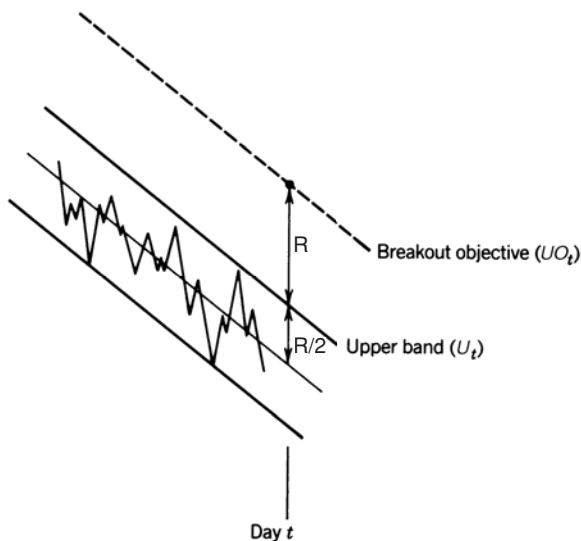
FIGURE 4.9   Channel calculation.

$$\text{Long objective (upper band), } UO_{t+1} = U_{t+1} + f \times \frac{R_t}{2}$$

$$\text{Short objective (lower band), } LO_{t+1} = L_{t+1} - f \times \frac{R_t}{2}$$

The objective may remain fixed at the price level determined on the day of the breakout, or preferably, will change each day to remain one band width from the new channel value (Figure 4.9). This method takes advantage of changing price volatility. More examples can be found in Chapter 20.

An alternate way of defining a channel would be to forecast one day ahead using the slope of a regression analysis (*linear* for a straight channel, *log* for a curved one) and use the standard deviation of the price changes to define the band. The other rules would remain the same.[12]

COMMODITY CHANNEL INDEX

The *Commodity Channel Index* (CCI) isn't necessarily for commodities and uses a channel only in the broadest sense. Instead, it is a measure of the deviation of the current price from the previous n days. It is best for mean reversion trading and has reasonable

---

[12] For a further discussion of channels, see Donald Lambert, "Commodity Channel Index," *Technical Analysis of Stocks & Commodities* (October 1980); and John F. Ehlers, "Trading Channels," *Technical Analysis of Stocks & Commodities* (April 1986).

statistical properties. First, find the average of the sum of the daily high, low, and close over n days (call it *ADP*),

$$ADP_t = \frac{\sum_{i=t-n+1}^{t} (H_i + L_i + C_i)}{n}$$

Then calculate the average deviation (AvgDev) over the same n-day interval

$$AvgDev_t = \frac{\sum_{i=t-n+1}^{i=t} |H_i + L_i + C_i - ADP_{i-1}|}{n}$$

Then CCI_t is the ratio of today's deviation divided by a fraction of the average deviation

$$CCI_t = \frac{(H_t + L_t + C_t - ADP_{t-1})}{0.015 \times AvgDev_{t-1}}$$

This can be easily done in a spreadsheet and plotted with the underlying price. The full spreadsheet, *TSM Commodity Channel Index HPQ*, can be found on the Companion Website. Development software may have a built-in function for this method.

The CCI is essentially a variation on a standard deviation channel with similar advantages and disadvantages. When prices become overbought during a strong upwards move, they can stay that way for weeks at a time. Simple rules for buying and selling oversold and overbought prices will give frequent small profits and an occasional very large loss.

WYCKOFF'S COMBINED TECHNIQUES

Richard D. Wyckoff, popular in the early 1930s and still discussed today, relied solely on charts to determine the motives behind price behavior. He combined the three most popular methods—bar charting, point charts (the predecessor of point-and-figure charts), and waves—to identify the direction, the extent, and the timing of price behavior, respectively.[13]

To Wyckoff, the bar chart combined price and volume to show the direction of the price movement. In general terms, it shows the trading ranges in which supply and demand are balanced. The volume complemented this by giving the intensity of trading, which relates to the *quality* of the long or short position. Wyckoff used *group charts*, or indexes, in the manner of Charles Dow, to select sets of stocks with the most potential, rather than looking only at individual stock price movement. This assures that the

---

[13]Jack K. Hutson, "Elements of Charting," *Technical Analysis of Stocks & Commodities* (March 1986).

move is based on the broader nature of the business, rather than on individual company dynamics. In today's market, we can use the S&P 500 or sector ETFs to accomplish the same objective.

Point-and-figure charts are used to condense price action. If prices move from lower to higher levels due to events, the time it takes to reach the new level is unimportant. Point-and-figure charts record events, not time. As long as prices rise without a significant reversal, the chart uses only one column; when prices change direction, a new column is posted (see the point-and-figure and the swing trading sections in Chapter 5). Price objectives can be determined from formations in a point-and-figure chart, and are usually related to the length of the sideways periods, or *horizontal formations*. These objectives are normally closer than objectives found using similar bar chart formations.

The *wave chart*, similar to Elliott's theories (discussed in Chapter 14), represents the behavior of investors and the natural rhythm of the market. Wyckoff uses these waves to determine the points of buying and selling within the limitations defined by both the bar chart and point-and-figure charts. He considered it essential to use the wave charts as a leading indicator of price movement.

Wyckoff used many technical tools but none rigidly. He did not believe in unconfirmed fundamentals but insisted that the market action was all you needed—the market's primary forces of supply and demand could be found in charts. He did not use triangles, flags, and other formations, which he considered to be a type of Rorschach test, but limited his analysis to the most basic patterns, favoring horizontal formations or congestion areas. He used time-based and event-based charts to find the direction and forecast price movement, then relied on human behavior (in the form of Elliott waves) for timing. His trading was successful, and his principles have survived.

COMPLEX PATTERNS

Most charting systems involve a few simple rules, trying to model a price pattern that seems to have repeatedly resulted in a profitable move. The most popular systems are trend breakouts, either a horizontal pattern or a trend channel. Over the years these approaches have proved to be steady performers. Another group of traders might argue that is it better to be more selective about each trade and increase the expectation of a larger profit than it is to trade frequently in order to win "in the long term"—that is, playing a statistical numbers game.

DeMark's Sequential™

Tom DeMark has created a strategy, called a *sequential*, that finds a very overextended price move, one that is likely to change direction, and takes a countertrend position.[14]

---

[14]Thomas R. DeMark, *The New Science of Technical Analysis* (New York: John Wiley & Sons, 1994).

His selling objective is to identify the place where the last buyer has bought. His rules use counting and retracements rather than mathematical formulas or trendlines. To get a buy signal, the following three steps are applied to daily data:

1. *Setup.* To begin, there must be a decline of at least nine or more consecutive closes that are lower than the corresponding closes four days earlier ($close_t < close_{t-4}$). In the case where today's close is equal to or greater than the close four days before, the setup must begin again.

2. *Intersection.* To assure that prices are declining in an orderly fashion, rather than plunging, the high of any day on or after the eighth day of the setup must be greater than the low of any day three or more days earlier. Note that there can be a delay before the intersection occurs provided that the pattern is not negated by the rules in Step 3.

3. *Countdown.* Once the setup and intersection have been satisfied, we count the number of days in which the close was lower than the close two days ago ($close_t < close_{t-2}$). The days that satisfy this countdown requirement do not need to be continuous. When the countdown reaches 13, we get a buy signal *unless* one of the following conditions occurs:

   a. There is a close that exceeds the highest intraday high that occurred during the setup stage.

   b. A sell setup occurs (nine consecutive closes above the corresponding closes four days earlier).

   c. Another buy setup occurs before the buy countdown is complete. In this case the rules begin again at Step 2. This condition is called *recycling*.

A sequential buy signal is shown in Figure 4.10 for the Deutsche mark (now the euro). The sell signal is the reverse of the buy. Traders should expect that the development of the entire formation will take no less than 21 days, but typically 24 to 39 days.

Entering the Sequential

Once the buy signal occurs, there are three choices for entering the market. The first is to enter on the close of the day on which the countdown is completed; however, this risks a new setup situation, which will extend the conditions for an entry. The second requires a confirmation of price direction, the close greater than the close four days ago, but it avoids the possibility of recycling. The third is to enter a long when the close is greater than the high two days earlier, a compromise between the first two techniques.

Exiting the Sequential

A number of exit conditions, consistent with the type of pattern, provide the trader with clear rules to liquidate the current trade. First, the current buy setup is complete, and the lowest price recorded does not exceed the furthest price recorded by the recent inactive setup (normally the previous sell setup). If, however, any price recorded in the current
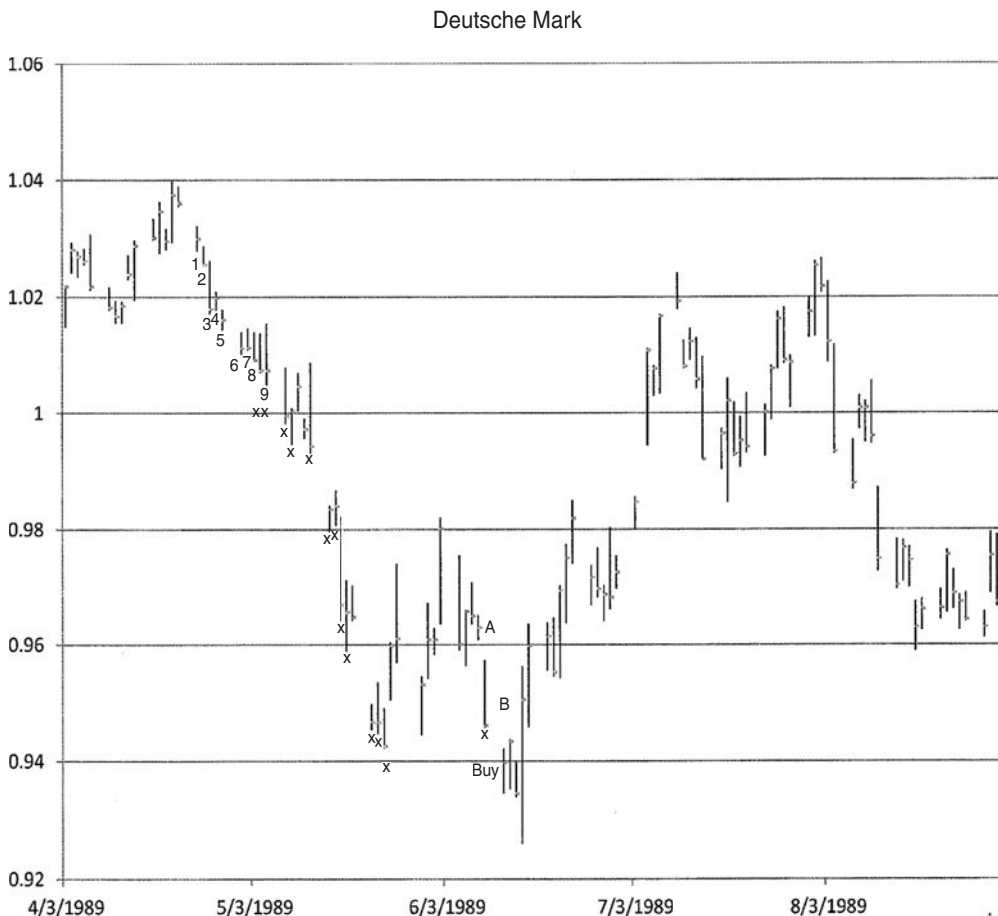
Deutsche Mark

FIGURE 4.10 A sequential buy signal in the Deutsche mark.

buy setup exceeds the furthest price of the previous sell setup, then the position is held until a reverse signal occurs.

Two stop-losses are also recommended. For a buy signal, the true range of the lowest range day of the combined setup and countdown period is subtracted from the low of that lowest day to create a stop-loss. Alternately, the difference between the close and the low of the lowest day is subtracted from the low of the lowest day for a closer stop-loss.

Thinking about Complex Patterns

There seems to be an extreme contrast between the simplicity of a trend breakout and the very complex set of circumstances that produce a signal for DeMark's sequential. The basic breakout system can be tested for robustness by comparing the performance of slightly longer and shorter calculation periods. As the calculation period becomes larger,

the trades should become more selective, the profits per trade become larger, and the overall performance profile improve. In the case of a single pattern, such as an island reversal or DeMark's sequential, there is no way to measure robustness in the same terms. For the sequential, there is only one count of 13 days and 1 pattern. Robustness can only be found by applying this method across different, unrelated markets. Each trader must decide whether this pattern, or any other complex set of rules, produces a better set of trades, or whether it is too specific to survive the test of time.

A STUDY OF CHARTING PATTERNS

There has always been a chasm between academics and practitioners over the value of technical analysis. Financial scholars have referred to charting as "voodoo finance," while technicians believe that their trading success is enough to refute academic conclusions that charting has little value. After all, it took many years of profitable public performance before trend following gained the attention of academics; there is no reason to expect something more complicated to take less time.

Computer Recognition of Chart Patterns

A very credible attempt to quantify charting patterns and assess their value was published by Lo, Mamaysky, and Wang.[15] The authors applied kernel regression as a smoothing technique, then defined 10 charting formations in the context of the smoothed price series. For example, a head-and-shoulders top formation is defined in terms of the most recent five local maxima and minima in the smoothed series, E_1, E_2, E_3, E_4, and E_5. In the definitions of the tops which follow, E_1, E_3, and E_5 are *maxima* and E_2 and E_3 are *minima*; for the bottom formations, which are not shown, E_1, E_3, and E_5 are *minima* and E_2 and E_3 are *maxima*.

Head-and-Shoulders Top

$E_3 > E_1$, $E_3 > E_5$
E_1 and E_5 are within 1.5% of their average
E_2 and E_4 are within 1.5% of their average

Broadening Top

$E_1 < E_3 < E_5$
$E_2 > E_4$

---

[15] Andrew W. Lo, Harry Mamaysky, and Jiang Wang, "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation," *Journal of Finance* (August 2000).

Triangular Top

$$E_1 > E_3 > E_5$$
$$E_2 < E_4$$

Rectangular Top

Tops are within 0.75% of their average
Bottoms are within 0.75% of their average
Lowest top > highest bottom

Double Top

E_1 and E_b are within 1.5% of their average, where E_1 is a maxima
$$t_a^* - t_1^* > 22$$

The identification of formations used a rolling window of 38 trading days; the notation t_1 represents the first day of the current window, 37 days back. We interpret the notation $t_a^* - t_1^* > 22$ to mean that the two extrema E_1 and E_b must be separated by more than 22 days.

Although the definitions are logical, the authors accept the differences between a mathematical definition of a charting formation and the visual, cognitive approach taken by a technical analysis. The human brain can assimilate and recognize more complex and subtle formations than the simple definitions presented in the paper. Then, on the one hand we have a somewhat limiting definition of chart patterns, and on the other we have the way in which humans select which patterns they choose to trade. It is far from certain which approach will yield the best returns.

The 10 patterns defined were head-and-shoulders tops and bottoms, broadening tops and bottoms, triangular tops and bottoms, rectangular tops and bottoms, and double tops and bottoms. In an example of the triangular formations, the key points used to identify the pattern did not align to form classic straight line sides; however, the consolidating formation that was recognized is itself a good candidate for analysis.

The success of the formation was measured by the returns over the three days immediately following identification. In addition, the formations were conditioned on the trend of volume; that is, returns were separated into formations that develop with increasing volume and decreasing volume.

Results of the Study

Tests were performed on several hundred U.S. stocks traded on both the NYSE and NASDAQ, from 1962 through 1996. The most common formations, the double top and double bottom, showed more than 2,000 occurrences of each. The next most frequent were the head-and-shoulders top and bottom, with over 1,600 appearances each. As a

control, a random, synthetically created price series was also tested and showed only ⅓ the number of head-and-shoulder formations. It argues that charting patterns are formed by the actions of the participants rather than by random events. Based on the number of stocks tested, the head-and-shoulders formation appeared about once each year for each stock.

All but one of the chart formations (the triangular top) showed positive returns for the three days following the identification of the formation. Of these, five were rated as *statistically significant*: the head-and-shoulders top, the broadening bottom, the rectangular top, the rectangular bottom, and the double top.

When formations were conditioned on rising or declining volume, the results changed for some of the patterns. In general, rising volume improved results. Falling volume was better for the head-and-shoulders top, and the rectangular top and bottom. Most analysts would expect rising volume to favor bottom formations and declining volume to improve most top formations; an academic study that contests this concept is likely to be viewed skeptically.

On the whole, technical analysts would not be disappointed with the conclusions of this study. Although the chart patterns may not meet the strict definition set by an experienced technician, they did capture the spirit of the formation and showed that positive returns followed. Confirmation is gratifying; any other conclusion would have been ignored.

For those not as mathematically gifted, but adept at computer programming, many of these formations can be created using the highs and lows generated from a swing chart, which is easily automated (a program is provided in Chapter 2). A trendline can be found using a least-squares regression through a series of swing highs or lows, qualified by a minimum variance. A triangle would be alternating swing highs and lows that are closer together. It is all within our reach.

BULKOWSKI'S CHART PATTERN RANKINGS

A particularly helpful section in Bulkowski's *Encyclopedia of Chart Patterns* is the summary at the end, where he ranks all the chart patterns by their success in forecasting price moves. The best five bullish formations are:

1. Upward breakout of a rectangular top
2. Upward breakout from a falling wedge
3. Upward breakout from an ascending triangle
4. Upward breakout from a double bottom
5. Upward breakout of a symmetric triangle after a downward move

For clarity, a rectangular top has multiple tests of a resistance level before the breakout, and a double bottom breaks out when the price goes above the highest price

that occurred between the two lows making up the double bottom. The upward breakout of a symmetric triangle formed during a downtrend is typically down and confirms the existing trend. Bulkowski's study shows that it performs the opposite way, and with high reliability.

The best five bearish patterns are

1. Descending scallops
2. Downward breakout of a symmetric triangle in a downtrend
3. Downward breakout of a broadening top
4. Downward breakout of a right-angled, descending, broadening formation
5. Downward breakout of a broadening bottom

Unlike the bullish formation, these seem dominated by broadening formations. For those unfamiliar with the *descending scallop*, an obscure formation, the *scallop* looks like a fishhook with the long stem to the left and the hook to the right. Attached to the end of the hook is another hook, also facing to the right, so that there is a series of longer declines and a shorter rounded recovery before another longer decline and rounded recovery.

A broadening top shows alternating swings getting larger, a sideways pause during an uptrend. With a right-angled, descending, broadening formation there is a somewhat horizontal top (the right-angle) to the swing highs and a broadening bottom before the downward break. A broadening bottom is similar to a broadening top but occurs during a downward price trend.

It is interesting that the broadening formations are more reliable when resolved to the downside. On the other hand, the symmetric triangle is the most dependable chart pattern because it forecasts a reliable price move whether it breaks to the upside or downside.

Event-Driven Trends

T rends are sustained price moves in one direction. In the previous chapter, an upwards trend was identified by drawing an upwards angling line under the lows of a rising price formation. But that involves a degree of interpretation. There are many other methods for recognizing the trend that are fully systematic. Of these there are two main categories, those that are sensitive to the way price moves over time, and others that ignore time and only concentrate on price level. This chapter will look at the second group.

A price trend can be thought of as an accumulation of small and large reactions to news and events. Even in a strong upwards trend, not all of the reactions are positive, but the net effect is a sustained upwards move. Smaller events that cause price movement are the result of frequent, scheduled economic reports, earnings reports by NYSE listed companies, upgrades and downgrades announced by financial institutions, and countless other pieces of information available on the various news media. Larger price moves come in reaction to natural disasters, political coups, unexpected election results, extreme action by a Central Bank, a major production interruption, terrorism, and unexpected corporate or political scandals of far-reaching effect. Each year there seem to be more events that lead to uncertainty and reduced investor confidence.

The impact of these material events can be long-term, temporary, or even structural, but they always cause an immediate change in price. Everyday news is filled with reasons for investors to either buy or sell, sometimes offsetting the effects of each other. Most often the result on price movement is small, but occasionally it can be a dramatic price shock.

Among the earliest of trading systems are those that signaled a new upwards trend when prices moved higher than they had been for some time. There is no math required, simply the idea that, if prices moved to a new high or new low, then something important has changed. This approach is intuitively sensible and has proved to be a successful strategy. As you progress through this book and become familiar with more complex and mathematically intricate techniques, continually ask yourself to what degree the newer methods have improved on the older, simpler ways of recognizing a trend.

We begin the presentation of trading systems with these event-driven trends that ignore time. A fundamental understanding of these methods is essential to every trader.

SWING TRADING

A *price swing* is a sustained price movement of a predetermined size. An upwards swing ends with a *swing high*, or peak, and a downward swing ends with a *swing low*, or valley. The distance from a peak to a valley is the swing. A swing can be small or large, depending on the sensitivity of the swing criterion. A swing is easily seen on a chart. You can choose to plot only the major price moves, or you can fill the chart with the smallest of price reversals. The only requirement is that each swing be greater than a threshold value, expressed in cents or dollars, or as a percentage of the current price of the stock or commodity. This minimum value, called the *swing filter*, determines the frequency of the swings and therefore the sensitivity of the chart.

Figure 5.1 shows a bar chart of gold futures with points above and below the bars representing the swing highs and lows based on a 5% minimum move from highs to low. Figure 5.2 shows the corresponding swing chart. Rather than 7 months of daily bars, there are only 18 points, each a new high or low separated from the previous by 5%. While the distance between these highs and lows can vary considerably on the daily bar chart, they have no time value on the swing chart. As long as prices do not make a new high or low by the swing filter amount, no entry occurs.

Constructing a Classic Swing Chart

A classic technique for finding the swing highs and lows uses the following seven steps[1] as shown in Figure 5.3. This can be applied to any data frequency, daily, weekly, hourly,
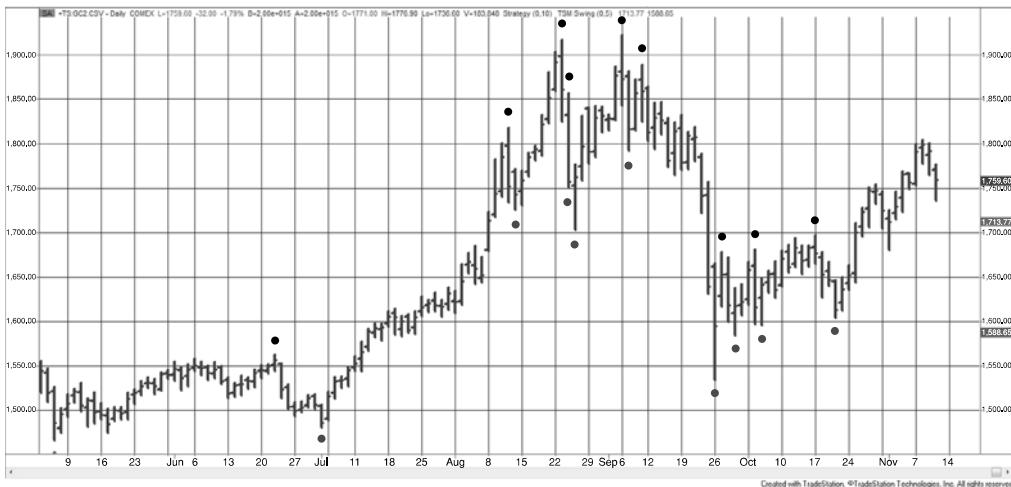
FIGURE 5.1   A bar chart of gold futures with 5% swing points marked.

[1]Based on William F. Eng, "A Mechanical Trading System," *Technical Analysis of Stocks & Commodities* (July 1986). Further information can be found in Eng, *The Technical Analysis of Stocks, Options & Futures* (Chicago: Probus, 1988).
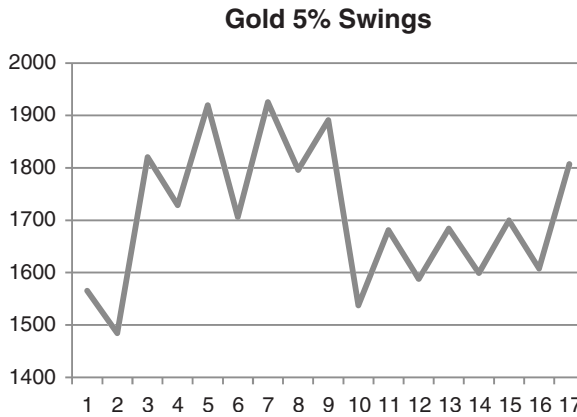
Gold 5% Swings

FIGURE 5.2   Corresponding swing chart of gold using a 5% swing filter.

or any intraday period; therefore, references to *days* in the following example could be replaced by *weeks* or *15-minute bars*. In the example below, we use daily data for gold cash prices and a swing threshold of $5 per ounce (not 5% as in the previous illustration).

1. Begin by plotting the high to low of the first day in the first column of the chart. At this point, we do not know the direction of this swing bar or the trend of gold prices.
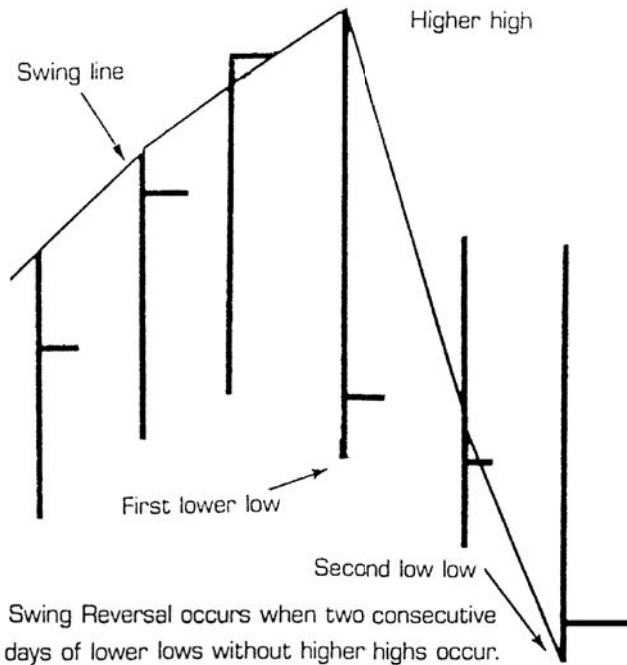
FIGURE 5.3   Constructing a classic swing chart.

2. Add the second day of data.

    a. If there is a higher high but not a lower low, then continue the first swing higher and ignore the low of the day. We can now call this an *upswing*.

    b. If there is a lower low but not a higher high, then continue the first swing lower and ignore the high of the day. This is now a *downswing*.

    c. If there is both a new high and a new low (an outside day), then continue the current high direction and ignore the low. As we will see later in the explanation of point-and-figure, you may choose the low instead, but this must choice must be consistently applied.

    d. If the high of the second day is less than the previous high of the column and the low is greater than the previous low (an inside day), then ignore that day and go to the next day.

3. Once a direction has been established, add the next day's data.

    a. If the current swing is up, then

        i. First look at the new daily high. If the high is higher than the high in the current column, extend that swing to the new high, ignoring the low. Go to the next day.

        ii. If the new high is lower than the high of the current column then test for a swing reversal. If the current swing high minus the low of the new day is greater than the swing filter value, then move one column to the right and plot the range of the new day high to low. The new swing is down.

    b. If the current swing is down, then

        i. Look first at the new daily low. If the low is lower than the low in the current column, extend that swing down to the new low and ignore the high. Go to the next day.

        ii. If the new low is higher than the low of the current column, then test for a swing reversal. If the high of the new day minus the low of the current column is greater than the swing filter value, then move one column to the right and plot the range of the new day from high to low. The swing is now up.

4. A new price that does not make a new high or low and does not exceed the swing reversal threshold is ignored.

This particular method, once the most popular, is called a *2-day* or *2-period* technique because it requires only two periods to identify a swing change. The sensitivity of the swing chart can be reduced by requiring that a swing change only occur after three or more days where prices reverse their direction.

CONSTRUCTING A SWING CHART USING A SWING FILTER

A very popular way of creating a swing chart uses a swing filter to define the sensitivity of the swings. The swing filter can be a value per share or a percentage of the current or

average share price. The following example uses cents per share. The basic eight rules for creating this swing chart are:

1. Select a stock or futures market and decide your chart sensitivity. For a stock with an average price of about $18, we select a swing filter of 75¢, about 4%.

2. Plot the first bar or day from the high to the low as a vertical line in the first column of the swing chart. The top of the bar is now called the *swing high* and the bottom is the *swing low*. Assume that prices are in an upswing. If the assumption is wrong, the chart will correct itself soon.

3. Move to the next day or, if intraday, move to the next bar.

4. If prices are in an upswing, continue with Step 5, otherwise go to Step 7.

5. Prices are in an upswing, and the high of the new day is higher than the current swing high. Extend the line higher, staying in the same column, to the point of the new high. Because prices are in an upswing, the lows of the bar are ignored.

6. Prices are in an upswing, and the high of the current bar is not higher than the current swing high. Look at the low price. If the swing high minus the current low is less than the swing filter of 75¢, then ignore this new price data. If the swing high minus the current bar low is equal to or greater than the swing filter, then move one column to the right and draw a new vertical line starting at the previous swing high and extending down to the current bar low. Connect the top of the previous bar to the top of the current bar. Continue with Step 3.

7. Prices are in a downswing, and the low of the new day is lower than the current swing low. Extend the line lower, staying in the same column, to the point of the new low. Because prices are in a downswing, ignore the highs of the bar.

8. Prices are in a downswing, and the low of the current bar is not lower than the current swing low, then look at the high. If the current high minus the swing low is less than the swing filter of 75¢, then ignore the new data. If the current bar high minus the swing low is equal to or greater than the swing filter, then move one column to the right and draw a new vertical line starting at the previous swing low and extending upwards to the current bar high. Connect the bottom of the previous bar to the bottom of the current bar. Continue with Step 3.

S&P Example

A good example that shows the construction of a swing chart can be seen using the S&P index, SPX, from October 1, 2010, through November 4, 2010, 25 days, shown in Figure 5.4. On the left in Table 5.1 are the high, low, and closing daily prices of the index, and on the right is the action taken. The gray areas represent the five swings identified during this period. The swing filter will be 20 big points.

Starting on day 1, the chart is initialized using the high and low of the day, 1150.30 and 1139.42. On day 2 there is no new high, and the low is 18.43 points below the previous
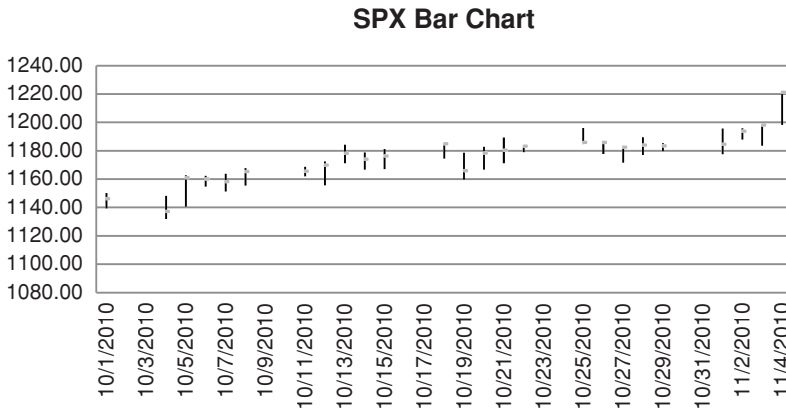
SPX Bar Chart

FIGURE 5.4 Classic bar chart showing SPX with gaps representing weekends (or any holiday).

TABLE 5.1 Day-by-Day Construction of a Swing Chart

	SPX Prices			Reversal of 20 points			
Date	**High**	**Low**	**Close**	**Day**	**Action**	**Price**	**Points**
10/1/2010	1150.30	1139.42	1146.24	1	Initialize		
10/4/2010	1148.16	1131.87	1137.03	2	Rev day		18.43
10/5/2010	1162.76	1140.68	1160.75	3	New high	1162.76	
10/6/2010	1162.33	1154.85	1159.97	4	Rev day		7.91
10/7/2010	1163.87	1151.41	1158.06	5	New high	1163.87	
10/8/2010	1167.73	1155.58	1165.15	6	New high	1167.87	
10/11/2010	1168.68	1162.02	1165.32	7	New high	1168.68	
10/12/2010	1172.58	1155.71	1169.77	8	New high	1172.58	
10/13/2010	1184.38	1171.32	1178.10	9	New high	1184.38	
10/14/2010	1178.89	1166.71	1173.81	10	Rev day		17.67
10/15/2010	1181.20	1167.12	1176.19	11	Inside day		
10/18/2010	1185.53	1174.55	1184.71	12	New high	1185.53	
10/19/2010	1178.64	1159.71	1165.90	13	Rev day		25.82
10/20/2010	1182.94	1166.74	1178.17	14	Rev day		23.23
10/21/2010	1189.43	1171.17	1180.26	15	New high	1189.43	
10/22/2010	1183.93	1178.99	1183.08	16	Rev day		10.44
10/25/2010	1196.14	1184.74	1185.62	17	New high	1196.14	
10/26/2010	1187.11	1177.72	1185.64	18	Rev day		18.42
10/27/2010	1183.84	1171.70	1182.45	19	Rev day		24.44
10/28/2010	1189.53	1177.10	1183.78	20	Rev up		17.83
10/29/2010	1185.46	1179.70	1183.26	21	Inside day		
11/1/2010	1195.81	1177.65	1184.38	22	Rev up		24.11
11/2/2010	1195.88	1187.86	1193.57	23	New high	1195.88	
11/3/2010	1198.30	1183.56	1197.96	24	New high	1198.30	
11/4/2010	1221.25	1198.34	1221.06	25	New high	1221.25	
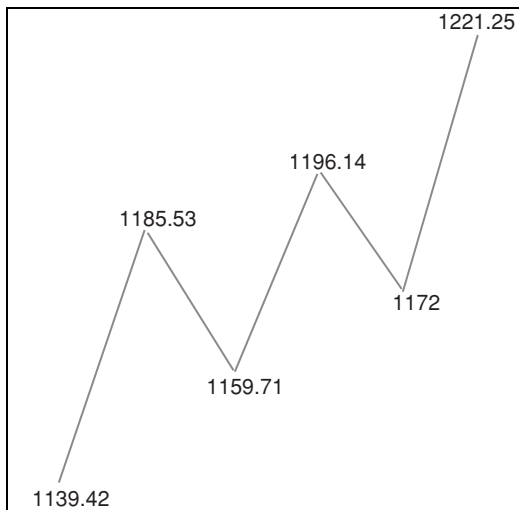
FIGURE 5.5 Swing chart of SPX.

high; therefore, the reversal criterion is not reached, and we continue with the next day. On day 3 there is a new high that establishes an upswing. On day 4 there is no new high, and the low indicates a reversal from the high of 7.91 points, far less than the amount needed to change the swing direction. This process is continued until day 13.

On day 13, the swing high is 1185.53, and the new low is 1159.71, 25.82 points, triggering a new downswing. On a chart, we move one column to the right, as seen in Figure 5.5. On day 14, there is another reversal of 23.23 points from the low of the previous day to the high of day 14, and large enough to change to an upswing. Then the first swing took 12 days to complete, and the second swing lasted only one day.

The swing chart corresponding to the SPX bar chart is shown in Figure 5.5. The extreme values of each swing are shown at the top and bottom of each swing. Note that there are only five columns in the swing chart, compared to 25 days in the bar chart. The swing chart ignores time; instead, it gives a schematic view of price movement, where the only relevant information are new swing highs and lows.

A different view of the swing chart construction is given in Figure 5.6. In this chart, the boxes are filled with the date on which the price moved into that box. It adds a little more information. Using this form of recording swing moves, it is possible to see the origins of point-and-figure charting, which will be explained in the next section.

Percentage Swings

The swing filter, which determines the swing high and low points, can be more robust if it is expressed as a percentage of price rather than as a fixed dollar per share or point value. Many markets have doubled in value—or halved—or both—over the past 10 years.

Price	Price swings				
	Up	Down	Up	Down	Up
1220					25
1215					25
1210					25
1205					25
1200					25
1195			17	19	22
1190			17	19	22
1185	12	13	15	19	22
1180	8	13	14	19	22
1175	8	13	14	19	22
1170	8	13	14		
1165	6	13	14		
1160	3	13	14		
1155	3				
1150	1				
1145	1				
1140	1				
1135	1				
1130					

FIGURE 5.6   Recording swings by putting the dates in the first box penetrated by the price.

During quiet market periods, stocks will split, causing the new share price, and the volatility, to be significantly lower than the day before. Using a fixed value for finding the swing highs and lows will cause the swing chart to be insensitive to price movement at low prices and to show frequent changes in swings at higher prices. The swing filter, expressed as a percentage p, avoids this problem.

$$\textit{Minimum swing value} \quad \text{MSV}_t = p \times \text{price}_t$$

This variable swing filter helps to keep the sensitivity of the swings the same over a long period, which is very helpful for backtesting of results and for more consistent trading signals. The only complication is that the minimum swing value may change daily, although gradually, and that requires you to be aware of the new value. It is always best to use the previous day's value, MSV_{t-1}, rather than today's value, in order to be sure that you are not cheating by looking ahead one day.

Rules for Swing Trading

Each vertical line on a swing chart, that is, each swing, represents a price trend. There are two sets of rules commonly used that enter positions in the trend direction:

1. (Conservative) *Buy* when the high of the current upswing exceeds the high of the previous upswing. *Sell* when the low of the current downswing falls below the low of the previous downswing.

2. (Active) *Buy* as soon as a new upswing is recognized. *Sell* when a new downswing is recognized. Both of these occur the first time there is a reversal greater than the swing filter.

An example of a buy signal can be seen in Figure 5.5, the SPX swing chart. The first swing peaked at 1185.53 before reversing to a short-term low of 1159.71. The next upswing moved above the old high; therefore, a buy order would have been placed at 1185.54, and a long position would have been entered. The end of the chart shows a high of 1221.25 which followed the swing low of 1172. If prices were to start down and penetrate that low, the long position would be exited.

The Swing Philosophy

The primary advantage of a swing method, and most event-driven trend systems, is that no action occurs if prices move sideways. We will see in the next two chapters that a trend recalculated on each bar, such as a moving average, has an agenda. That is, prices must continue to advance if the trend is to remain intact. In a swing philosophy, prices can move sideways or stand still within a trend. Prices can move up and down in any pattern as long as they do not violate the previous swing highs (if in a downtrend) or swing lows (if in an uptrend). This characteristic of event-driven systems makes them very robust *at a cost of higher risk*. Risk will be measured as the difference between the entry point of a trade (the price at which the old swing high or low was penetrated) and the price at which a reverse trade would be entered. This risk can be as small as a swing reversal or much larger if prices move quickly without any intermediate reversals.

A significant benefit of the swing method is that it can signal a new trade at the moment of a significant event. If the news is a surprise to the market and prices move out of their current trading levels to new highs or new lows, the swing method will signal a new trade. This immediate response is a very positive feature for most traders, who want to act in a timely manner. It is very different from trend systems that use moving average or other *time series calculations*, which have lags.

The following systems are simple variations of the swing method of charting.

The Livermore System

Known as the greatest trader on Wall Street, Jesse Livermore was associated with every major move in both stocks and commodities during the 30-year period from 1910 to 1940. Livermore began his career as a board boy, marking prices on the high slate boards that surrounded the New York Stock Exchange floor. During this time, he began to notice the distinct patterns in the price movement that appeared in the columns of numbers.[2] As

---

[2]Edwin Lefèvre, *Reminiscences of a Stock Operator* (Burlington, VT: Books of Wall Street, 1980). First published by George H. Doran, 1923.
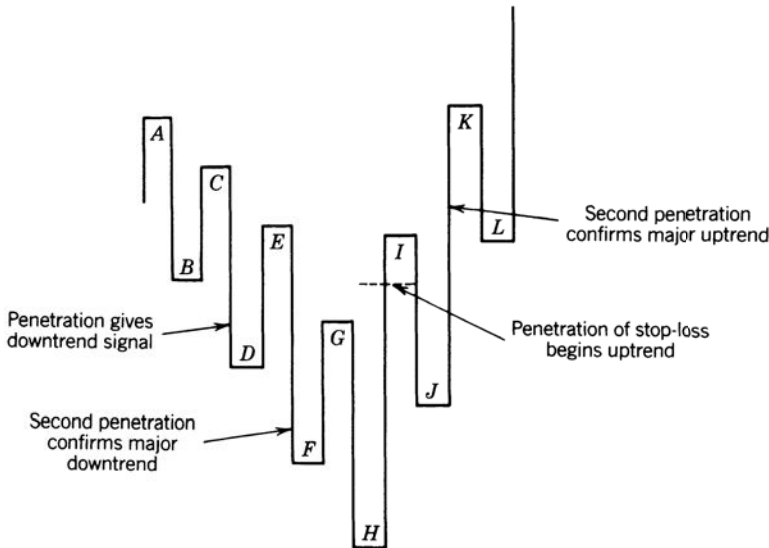
FIGURE 5.7 Livermore's trend change rules.

Livermore developed his trading skills and eventually took his position as a professional trader, he maintained the habit of writing prices in columns headed *Secondary Rally*, *Natural Rally*, *Up Trend*, *Down Trend*, *Natural Reaction*, and *Secondary Reaction*. This may have been the basis for what is now a swing chart.

Livermore's approach to swing trading required two filters, a larger *swing filter* and a *penetration filter* of one-half the size of the swing filter. Penetrations were significant at price levels he called *pivot points*. A pivot point is defined in retrospect as the top and bottom of each new swing; the pivot points are marked with letters in Figure 5.7.[3] Pivot points remain a popular way of identifying relative highs and lows.

Livermore's trading technique is a unique interpretation of the swing chart. Positions are taken only in the direction of the major trend. A *major uptrend* is defined by confirming higher highs and higher lows, and a *major downtrend* by lower lows and lower highs, and where the *penetration filter* (swing filter) is not broken in the reverse direction. That is, an uptrend is still intact as long as prices do not decline below the previous pivot point by as much as the amount of the penetration filter (seen in Figure 5.7). Once the trend is identified, positions are added each time a new penetration occurs, confirming the trend direction. A stop-loss is placed at the point of penetration beyond the prior pivot point. Unfortunately, Livermore never revealed how the penetration point was calculated. It seems, however, to be a percentage (for example, 20%) of the current swing size.

---

[3]Jesse Thompson, "The Livermore System," *Technical Analysis of Stocks & Commodities* (May 1983).
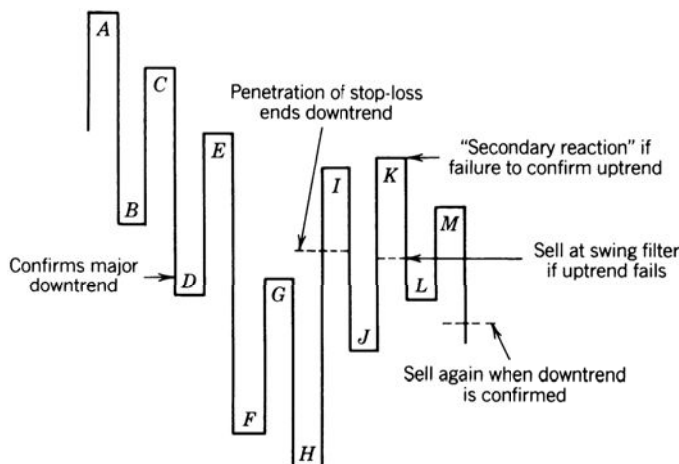
FIGURE 5.8   Failed reversal in the Livermore method.

Failed Reversal

In the Livermore system, the first penetration of the stop-loss (a swing high or low depending on the direction of the trade) calls for liquidation of the current position. A second penetration is necessary to confirm the new trend. If the second penetration fails (at point *K* in Figure 5.8), it is considered a secondary reaction within the old trend. The downtrend may be reentered at a distance of the swing filter below *K*, guaranteeing that point *K* is defined, and again on the next swing, following pivot point *M*, when prices reach the penetration level below pivot point *L*. It is easier to reenter an old trend than to establish a position in a new one.

Programming the Swing High and Low Points

The swing method can be programmed in strategy development software or Excel. The good part about specialized software is that the points can be shown on a price chart to give visual confirmation, as seen in Figure 5.1. The indicator *TSM Swing* can be found on the Companion Website along with the Excel spreadsheet that does the same calculations.

Keltner's Minor Trend Rule

One of the classic trading systems is the *Minor Trend Rule* published by Keltner in his book, *How to Make Money in Commodities*. For many years it was followed closely by a great part of the agricultural community and should be understood for its simplicity and potential impact on markets. In today's high-tech environment, it is important to remember that many trading decision are still made using simple tools.

Keltner defines an upward trend by the failure to make new lows (comparing today's low with the prior day) and a downtrend by the absence of new highs. This notion is consistent with chart interpretation of trendlines by measuring upward moves along the bottom and downward moves along the tops. The Minor Trend Rule is a plan for using the daily trend as a trading guide. The rule states that the minor trend turns up when the daily trend trades above its most recent high; the minor trend stays up until the daily trend trades below its most recent low, when it is considered to have turned down. In order to trade using the Minor Trend Rule, *buy* when the minor trend turns up and *sell* when the minor trend turns down; always reverse the position.

The Minor Trend Rule is a simple short-term trading tool, buying on new highs and selling on new lows with risk varying according to volatility. It is a breakout method in the style of swing trading and is the basis for a number of current technical systems that vary the time period over which prior highs and lows are established and consequently increase the interval between trades and the risk of each trade. An advantage of the Keltner approach is that, as with swing trading, it imposes no arbitrary restrictions on the size of the price move, such as a breakout of 100 points.

Pivot Points

A *pivot point* is the high or low point of a reversal, but a much weaker condition than a swing point. It could be the center point of 3 trading days, or 5 days, or more. A 5-day pivot point means that there are 2 days on each side of the local high or low price. It also means that a 5-day pivot point has a 3-day lag because you cannot identify it until the close of the third day (the second day after the high or low point). Pivot points will be used in strategies throughout this book.

Wilder's Swing Index

Although Wilder called this the *Swing Index*, it is a combination of daily range measurements with trading signals generated using *pivot points*. However, it is clearly an event-driven method. It was presented with trading rules in Wilder's *Swing Index System*.[4] Wilder determined that the five most important positive patterns in an uptrend are:

1. Today's close is higher than the prior close.
2. Today's close is higher than today's open.
3. Today's high is greater than the prior close.
4. Today's low is greater than the prior close.
5. The prior close was above the prior open.

---

[4]J. Welles Wilder, Jr., *New Concepts in Technical Trading Systems* (Greensboro, NC: Trend Research, 1978).

In a downtrend, these patterns are reversed.

The Swing Index, *SI*, combines these five factors, then scales the resulting value to fall between +1 and –1.

$$SI_t = 50 \times \left(\frac{(C_t - C_{t-1}) + 0.5(C_t - O_t) + 0.25(C_{t-1} - O_{t-1})}{TR_t} \right) \times \frac{K}{M}$$

where K = the largest of $\left| H_t - C_{t-1} \right|$ and $\left| L_t - C_{t-1} \right|$
M = the value of a limit move (*more about this below*)
TR = the *true range*

True range is calculated from the following two steps (note that this is the same true range as commonly used, however, step 2 requires that you know which of the three combinations was largest):

1. First, determine which is the largest of
   a. $\left| H_t - C_{t-1} \right|$
   b. $\left| L_t - C_{t-1} \right|$
   c. $\left| H_t - L_t \right|$
2. Calculate R according to the corresponding formula, using (a) if the largest in step 1 was (a), or using (b) or (c), depending on which one of those was the largest in step 1.
   a. $R = \left| H_t - C_{t-1} \right| - .5 \times \left| L_t - C_{t-1} \right| + .25 \times \left| C_{t-1} - O_{t-1} \right|$
   b. $R = \left| L_t - C_{t-1} \right| - .5 \times \left| H_t - C_{t-1} \right| + .25 \times \left| C_{t-1} - O_{t-1} \right|$
   c. $R = \left| H_t - L_t \right| + .25 \times \left| C_{t-1} - O_{t-1} \right|$

The *SI* calculation uses three price relationships: the net price direction (close-to-close), the strength of today's trading (open-to-close), and the memory of yesterday's strength (prior open-to-close). It then uses the additional factor of volatility as a percentage of the maximum possible move (*K/M*). The rest of the formula simply scales the results to within the range of +1 to –1.

Updating the Swing Index

In today's market, *M*, the *limit move*, is a problem. When Wilder developed this method, all markets had limit moves, that is, trading halted when prices moved to a maximum daily price change as determined by the exchange. That is no longer the case. With some markets, trading is temporarily halted when prices move to a preset limit, but after a few minutes trading begins again. For those markets, there is no clear value for *M*; however, that is easily resolved. The function of *M* is used only to scale the index into the range +1 to –1. If we choose an arbitrary value, larger than the normal daily move, the results will be fine. Because the trading rules use only the relative highs and lows of the index, rather
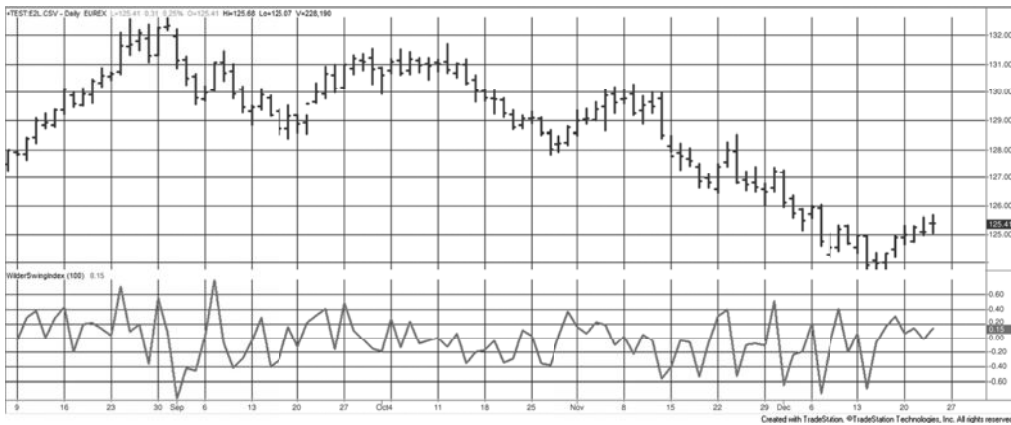
FIGURE 5.9   Wilder's Swing Index applied to Eurobund continuous futures, using TradeStation modified code.

than thresholds such as 0.90, the results will be the same, regardless of the choice of M. In these examples $M = 100$.

Figure 5.9 shows the daily prices for the Eurobund futures along with the values of the Swing Index for the last few months of 2010 and the beginning of 2011. *SI* switches from a positive bias at the beginning of the chart to a negative one as prices turn from bullish to bearish. These calculations are given in the spreadsheet, *TSM Wilder Swing Index*, found on the Companion Website. The Swing Index is also a built-in indicator and function in TradeStation; however, the code requires that the limit move be pre-set with each market. Instead, the program *TSM Wilder Swing Index*, provided on the Companion Website, allows you to input that value as, for example, 100.

Trading Rules

The daily SI_t values are added together to form an *Accumulated Swing Index* (*ASI*),

$$ASI_t = ASI_{t-1} + SI_t$$

which is substituted for the price and used to generate trading signals, allowing ready identification of the significant highs and lows as well as clear application of Wilder's trading rules. The terms used in the trading rules are:

> *HSP, High swing point*—Any day on which the *ASI* is higher than both the previous and the following day.
> *LSP, Low swing point*—Any day on which the *ASI* is lower than both the previous and the following day.
> *SAR, Stop and reverse points* (three types)—*Index SAR* points generated by the *ASI* calculation, *SAR* points applied to a specific price, and *Trailing Index SAR*, which lags 60 *ASI* points behind the best *ASI* value during a trade.

The Swing Index System rules are:

1. Initial entry:

   a. Enter a new long position when the ASI_t crosses above HSP_{t-2}.

   b. Enter a new short position when the ASI_t crosses below LSP_{t-2}.

   Note that HSP and LSP cannot be identified until two days after it occurs.

2. Setting the *SAR* point:

   a. On entering a new long trade, the *SAR* is the most recent *LSP*; the *SAR* is reset to the first *LSP* following each new *HSP*. A *trailing SAR* is determined as the lowest daily low occurring between the highest *HSP* and the close of the day on which the *ASI* dropped 60 points or more.

   b. On entering a new short trade, the *SAR* is the most recent *HSP*; the *SAR* is reset to the first *HSP* following each new *LSP*. The *trailing SAR* is determined as the highest daily high occurring between the lowest *LSP* and the close of the day on which the *ASI* rose by 60 points or more.

The program *TSM Accumulated Swing Index* can be found on the Companion Website.

POINT-AND-FIGURE CHARTING

There does not appear to be any record of which came first, swing charting or point-and-figure charting. Both methods are very similar; however, point-and-figure has developed a much more extensive following. Point-and-figure charting is credited to Charles Dow, who is said to have used it just prior to the turn of the twentieth century. It has three important characteristics:

1. It has simple, well-defined trading rules.
2. It ignores price reversals that are below a minimum price move as determined by the *box size*.
3. It has no time factor (it is event-driven). As long as prices fail to change direction by the reversal value, the trend is intact.

When point-and-figure charting first appeared, it did not contain the familiar boxes of *X*s and *O*s. The earliest book containing the subject is reported to be *The Game in Wall Street and How to Play it Successfully*, published by Hoyle (not Edmond Hoyle, the English writer) in 1898. The first definitive work on the subject was by Victor De Villiers, who in 1933 published *The Point and Figure Method of Anticipating Stock Price Movement*. De Villiers worked with Owen Taylor to publish and promote a weekly point-and-figure service, maintaining their own charts; he was impressed by the simple, scientific methodology. As with many of the original technical systems, the application

Silver

854.0	X								
853.0	X	O	X						
852.0	X	O	X	O					
851.0		O	X	O	X				
850.0		O		O	X	O			
849.0				O	X	O			
848.0				O	X	O			
847.0				O					
846.0									
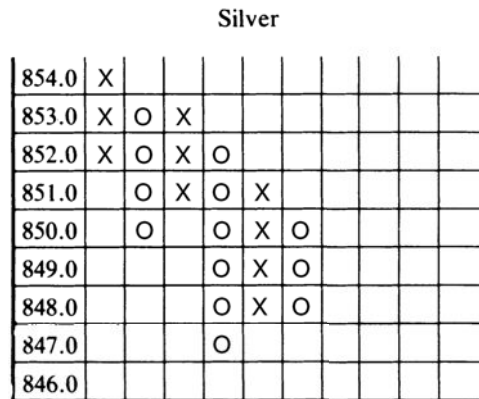
FIGURE 5.10   Point-and-figure chart.

was intended for the stock market, and the rules required the use of every price change appearing on the ticker. It has also been highly popular among futures traders in the grain and livestock pits of Chicago. The rationale for a purely technical system has been told many times, but an original source is often refreshing. De Villiers said:[5]

The Method takes for granted:

1. *That the price of a stock at any given time is its correct valuation up to the instant of purchase and sales (a) by the consensus of opinion of all buyers and sellers in the world and (b) by the verdict of all the forces governing the laws of supply and demand.*
2. *That the last price of a stock reflects or crystalizes* everything *known about or bearing on it from its first sale on the Exchange (or prior), up to that time.*
3. *That those who know more about it than the observer cannot conceal their future intentions regarding it. Their plans will be revealed in time by the stock's subsequent action.*

The unique aspect of the point-and-figure and swing methods is that they ignore the passage of time The point-and-figure chart differs from the swing chart in that each column representing an upswing is a series of boxes containing *X*s, and each downswing is shown as a string of *O*s (Figure 5.10), and a mark is not placed unless a minimum price change occurs.

The original *figure charts* were traditionally plotted on graph paper with square boxes, and only dots, or the exact price, were written in each box. The chart evolved to have prices written on the left scale of the paper, where each box represented a

---

minimum price move. Some point-and-figure chartists then used a combination of *X*s and occasional digits (usually 0s and 5s every five boxes) to help keep track of the length of a move. In some cases the top of an upswing column was connected to the start of a downswing in the next column with a crossbar, and the bottom of a downswing column was connected to the beginning of the next upswing column. This gave the point-and-figure chart an appearance similar to the swing chart. Charts using 1, 3, and 5 points per box were popular, where each point represented a minimum price move. In the 5-point method, no entry was recorded unless the price change spanned 5 points.

Point-and-figure charts, which were commonly used on the floor of the Chicago Mercantile Exchange and the Chicago Board of Trade up to the late 1990s, are intended to show the greatest detail. Each box represents the minimum allowable price move, and reversals of direction use the traditional 3-box criteria. Floor traders use the charts to show only the short-term price moves, and leave a lot to the interpretation of patterns.

Plotting Prices Using the Point-and-Figure Method

To plot prices on a point-and-figure chart, start with a piece of square-box graph paper and mark the left scale using a conveniently small price increment. For example, each box may be set at $0.25 for Microsoft, $0.50 for IBM, 5.0 points for S&P 500 index, $1 for gold and platinum, 4/32 for 30-year bond futures, 1¢ for soybeans and silver, and so forth (as in Figure 5.10). The choice of a box size will make the chart more or less sensitive to changes in price direction as will be seen in later examples. The smaller the box size, the more changes in direction will be seen. This also corresponds to longer and shorter trends or major and minor trends. Therefore, a point-and-figure chartist looking for a long-term price movement will use a larger box size. Box sizes are often related to the current volatility of the markets.

Once the graph paper has been scaled and the prices entered along the left side, the chartist can begin. The first box is entered with the current closing price of the market. If the price of silver is 852.50 and a 1¢ (1¢ = 1.00) box is being used, a mark is placed in the box beside the value 852. An *X* or an *O* is used to indicate that the current price trend is up or down, respectively. Either an *X* or *O* may be used to begin—after that, it will be determined by the method.

The rules for plotting point-and-figure charts are easily shown as a flowchart in Figure 5.11. Preference is given to price movements that continue in the direction of the current trend. Therefore, if the trend is up (represented by a column of *X*s), the new high price is tested first; if the trend is down, the low price is given preference. The opposite price is checked only if the new price fails to increase the length of the column in the direction of the current trend.

The traditional point-and-figure method calls for the use of a *3-box reversal*, that is, the price must reverse direction by an amount that fills 3 boxes from the most extreme box of the last column before a new column can begin (it actually must fill the fourth box because the extreme box is left blank). The importance of keeping the 3-box
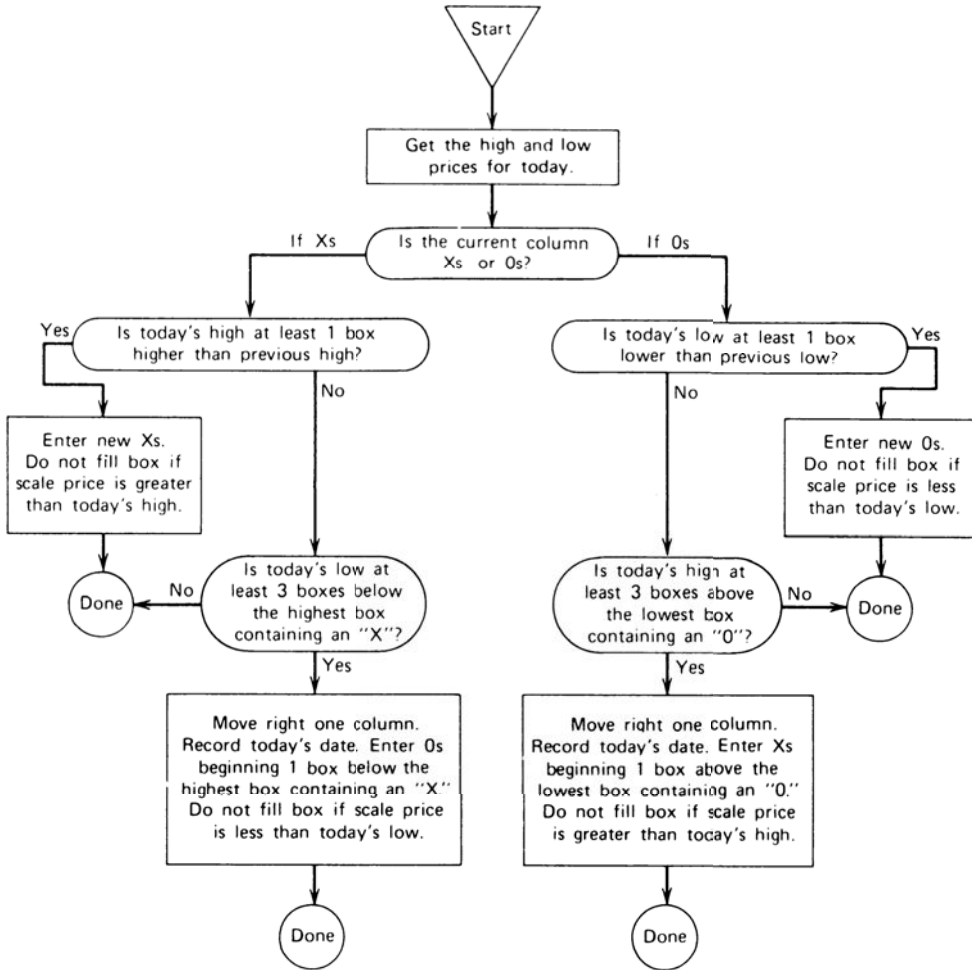
FIGURE 5.11   Point-and-figure daily rules.

reversal has long been questioned by experienced point-and-figure traders. It should be noted that the net reversal amount (the box size times the number of boxes in the reversal) is the critical value. For example, a 5-point box for the NASDAQ 100 ETF (QQQ) with a 3-box reversal means that QQQ prices must reverse from the lows of the current downtrend by 15 points to indicate that an uptrend has started. The opposite combination, a 3-point box and a 5-box reversal, would signal a new trend at the same time, after a 15-point reversal. The difference between the two choices is that the smaller box size would recognize a smaller continuation of a price move by filling more boxes. Ultimately, the smaller box size will capture more of the price move; it is considered the preferable alternative. The choice of box size and reversal boxes will be considered later in more detail.

Painless Point-and-Figure Charts

There are a number of graphic charting and quote systems that allow a simple bar chart to be converted to point-and-figure automatically. It is still necessary to specify the box size and the reversal size. The reason for showing the construction in detail is that none of these services provide trading signals or performance results based on point-and-figure charting. For that, it will be necessary to code the instructions into a spreadsheet or a strategy development platform.

Point-and-Figure Chart Formations

It would be impossible for the average speculator to follow the original method of recording every change in price. When applied to stocks, these charts became so lengthy and covered so much paper that they were unwieldy and made interpretation difficult. In 1965, Robert E. Davis published *Profit and Profitability*, a point-and-figure study that detailed eight unique buy and sell signals. The study covered two stocks for the years 1914–1964, and 1100 stocks for 1954–1964. The intention was to find specific bull and bear formations that were more reliable than others. The study concluded that the best buy signal was an *ascending triple top* and the best sell signal was the *breakout of a triple bottom*, both shown in Figure 5.12 and with the other patterns studied in Figure 5.13.

Plotted using daily data, futures prices do not offer the variety of formations available in the stock market. The small number of markets and the high correlation of movement between many of the index and interest rate markets make the limitations of signal selection impractical. Instead, the most basic approach is used, where a buy signal occurs when an *X* in the current column is one box above the highest *X* in the last column of *X*s, and the simple sell signal is an *O* plotted below the lowest *O* of the last descending column. The flexibility of the system lies in the size of the box; the smaller the size, the more sensitive the chart will be to price moves. In 1933, Wyckoff noted that it was advisable to use a chart with a different box size when the price of the stock varied substantially.[6]
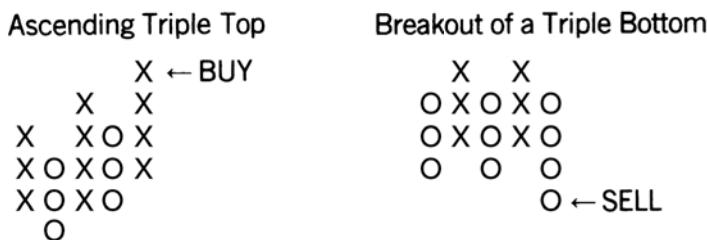
FIGURE 5.12   Best formations from Davis's study.

---

[6]Richard D. Wyckoff, *Stock Market Technique*, *Number One* (New York: Wyckoff, 1933), 89.

(a)

Breakout of a triple top

Ascending triple top

Spread triple top

Upside breakout of a
symmetric triangle

Upside breakout of a
bullish resistance line

Upside breakout of a
bearish resistance line

Upside breakout of an
ascending triangle

Bullish catapult

(b)

Breakout of a triple bottom

Descending triple bottom

Spread triple bottom

Downside breakout of a
symmetric triangle

Downside breakout of a
bearish support line

Downside breakout of a
bullish support line

Downside breakout of a
descending triangle

Bearish catapult

FIGURE 5.13　　(a) Compound point-and-figure buy signals. (b) Compound point-and-figure sell signals.
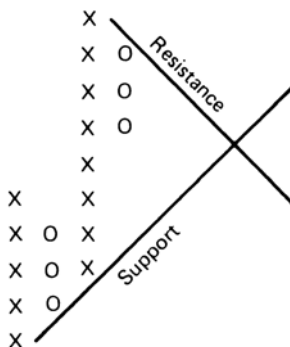
FIGURE 5.14 Point-and-figure trendlines.

Point-and-Figure Trendlines

Bullish and bearish trendlines are commonly used with point-and-figure charts. The top or bottom box that remains blank when a reversal occurs can form the beginning of a descending or ascending pattern at a 45° angle (diagonally through the corners of the boxes, providing the graph paper has square boxes). These 45° lines represent the major anticipated trends of the market. Once a top or bottom has been identified, a 45° line can be drawn down and to the right from the upper corner of the top boxes of *X*s, or up and towards the right from the bottom of the lowest box of *O*s (Figure 5.14). These trendlines are used to *confirm* the direction of price movement and are often used to filter the basic point-and-figure trading signals so that only long positions are taken when the 45° trendline is up and only shorts sales are entered when the trendline is down.

More Point-and-Figure Studies

In 1970, Charles C. Thiel, Jr., with Robert E. Davis, completed the first purely futures market point-and-figure study[7] that calculated profitability of a reasonably large sample of markets by varying both the value of a box and the reversal criteria. With the standard 3-box reversal and only simple buy and sell formations, the tests showed 799 signals, of which 53% were profitable; the average net profit on all trades was $311 realized in approximately 50 days. The period studied was 1960 through 1969. In the mid-1970s, Zieg and Kaufman[8] performed a computerized study using the same rules but limiting the test period to six months ending May 1974, an extremely active

---

[7]Charles Thiel and R. E. Davis, *Point and Figure Commodity Trading: A Computer Evaluation* (West Lafayette, IN: Dunn & Hargitt, 1970).

[8]Kermit C. Zieg, Jr., and Perry J. Kaufman, *Point and Figure Commodity Trading Techniques* (Larchmont, NY: Investors Intelligence, 1975). This book contains complete tabularized results of both point-and-figure tests.

market period. For the 22 commodities tested, 375 signals showed 40% of the trades were profitable; the net profit over all the trades was $306 and the average duration was 12.4 days. It is interesting to note that the most significant difference in the results of the two studies is in the average length of a trade, from 50 in the Thiel and Davis study to 12.4 days in the Zieg and Kaufman tests, indicating a change apparently induced by more volatile markets. Although the two tests varied in many of the details, the results are a strong argument for the consistency of the point-and-figure method as a trading tool.

In its current role, point-and-figure differs from traditional charting because it provides a rigid set of trading rules. Many of the formations are still subject to interpretation and are frequently used that way by floor traders. For the more systematic trader, it will tell exactly what penetration of a resistance or support level is necessary to generate a buy or sell signal and exactly where the stop-loss order should be placed to limit risk. It is this well-defined nature of point-and-figure charting that allows computer testing and evaluation.

A complete study of the point-and-figure method includes rules of charting, buy and sell signals, trendlines, geometric formations, and price objectives. A discussion of these features can be found in Chapter 3. They apply equally to point-and-figure charting. They have also been covered effectively in a book by Cohen and another by Zieg and Kaufman.[9] The following sections cover more advanced point-and-figure topics, including its relationship to bar charting, alternate plotting rules, risk-limited trading, and varying box size.

Point-and-Figure Box Size

The box size used in a point-and-figure chart determines the sensitivity, or frequency of signals. The selection of the box size is critical to successful trading. For many years Chartcraft (Investors Intelligence) was the only major service that produced a full set of point-and-figure charts for the futures markets. A history of their box sizes is shown in Table 5.2. There are now a number of services providing these charts, and they can easily be found by searching the Internet for "point-and-figure charts."

Since the 1970s, every traded commodity has had at least one major price move taking it to levels often greater than twice the normal price. Sugar and silver each topped at 10 times their value in 1970. By 2000 many technology stocks had surged 20 times their 1990 value only to retreat by as much as 90% in the next few years. These moves necessitate changes in box size in order to control the impact of the increased, and later decreased, volatility.

Table 5.3 shows the performance of the point-and-figure method from 2000 through 2010 for a selection of widely traded futures markets. These were created using the program *TSM Point & Figure*, available on the Companion Website. When testing these

---

[9]A. W. Cohen, *How To Use the Three-Point Reversal Method of Point and Figure Stock Market Trading* (Larchmont, NY: Chartcraft, 1972); and Kermit C. Zieg and Perry J. Kaufman, *Point & Figure Commodity Trading Techniques* (Larchmont, NY: Chartcraft, 1975).

TABLE 5.2 Point-and-Figure Box Sizes*

Futures Market	Units	Prior to 1975† Year	Prior to 1975† Box Size	1975‡ Box Size	1977‡ Box Size	1977§ Box Size	1986‡ Box Size	2002–2003 Box Size
Grains								
Corn	cents	1971	½	2	2	2	1	2
Oats	cents	1965	½	1	1		1	½
Soybeans	cents	1971	1	10	10	5	5	4¼
Soybean meal	pts	1964	50	500	500		100	200
Soybean oil	pts	1965	10	20	20		10	20
Wheat	cents	1964	1	2	2	2	1	1.5
Livestock and Meats								
Live cattle	pts	1967	20	20	20		20	10
Live hogs	pts	1968	20	20	20		20	40
Pork bellies	pts	1965	20	20	20		20	75
Other Agricultural Products								
Cocoa¶	pts	1964	20	100	100	50	10	2
Coffee	pts		(20)†	100	100	50	100	
Cotton	pts		(20)†	100	100		50	90
Lumber	pts		(100)†	100	100		100	170
Orange juice	pts	1968	20	20	100	20	100	100
Sugar	pts	1965	5	20	20		10	16
Metals								
Copper	pts	1964	20	100	100	50	50	100
Gold	pts			50	100		400	200
Platinum	pts	1968	200	100	200	200	400	
Silver	pts	1971	100	200	200	400	1000	150

*All box sizes use a 3-box reversal and are in points (decimal fractions treated as whole numbers) unless otherwise indicated.

†Cohen (1972); parentheses indicate approximate values.

‡Courtesy of Chartcraft Commodity Service, Chartcraft, Inc., Larchmont, New York.

§Chart Analysis Limited, Bishopgate, London. Values are for long-term continuation charts.

¶Cocoa contract changed from cents/pound to dollars/ton.

markets, the first problem that surfaces is that each market needs its unique box size. For example, soybeans trading at $15/bushel will need a box size smaller than the DAX, trading at 6000. The box size reflects a sensitively to volatility during the period of the test data.

To solve that problem as simply as possible, the box sizes were initialized to a percentage of the price at the beginning of the test. These centered around 1%, so that gold at $1000/oz would have a box size of $10. All gold signals used the $10 box size and a

TABLE 5.3a Net Profits for the Point-and-Figure Method
Initial box sizes range from 0.10% to 2.0% for a selection of futures markets, 2000–2010. Negative results are in grey, and the best results for each market are outlined.

Box (%)	Crude Oil	Cotton	DAX	EUR	Bund	Euro-dollars	S&P	Gold	Heating Oil	Japanese Yen	NASDAQ	Soy-beans	10-Year Notes	30-Year Bonds	Wheat
2.0	3540	-12540	31450	54538	0	0	-6838	-10200	-22142	-925	4895	-8725	19641	-5188	-26200
1.9	-10510	-22970	80800	74600	0	0	-11325	3450	-10580	14413	-2165	11150	16484	14813	-28813
1.8	-42370	-44990	186225	74875	14360	0	11975	-24610	-55797	-9563	-11730	650	34672	2906	-32288
1.7	-9840	-45610	133513	89763	7240	0	-11500	25300	-122653	14100	-13985	18275	13953	26156	-32750
1.6	-28630	-67010	42425	104875	14580	0	-6875	-740	-97297	-14363	-21930	9700	24766	-3438	-26688
1.5	-34170	-47705	-4738	88325	2840	0	12225	-2860	-82555	-31538	-1215	8650	8328	17156	-7663
1.4	26410	-42895	74125	89125	-9120	0	2400	12480	-60967	-1263	-21130	10900	-2953	49563	-13363
1.3	6860	-53445	123688	105038	-6350	0	-25850	16080	-60955	-11725	-29685	8000	-1766	20219	-13225
1.2	-13150	-53080	105950	128650	-9260	0	-20000	19980	-40110	-3488	-29150	-18000	16	2719	-1013
1.1	46420	-39220	92313	119838	710	0	-37300	33020	-23822	-5150	-43535	-5275	-1328	33688	2750
1.0	46420	-33535	175338	107063	20890	0	-41250	-15280	-57464	-9175	-24410	-12550	-9609	23313	-15113
0.9	87410	-6700	-16375	117800	31980	0	-46500	9820	-61828	-66838	-12890	-22025	1828	10875	2538
0.8	32520	-3920	-15975	111700	22820	0	1250	-13020	-48313	-57213	-350	-7000	2078	60688	-24475
0.7	30770	2310	-8888	90063	16360	0	-1363	2500	-56268	-90025	15010	1925	1016	52500	-14963
0.6	-27060	25630	888	47263	21830	0	-40763	120	-82870	-50163	-16020	-11325	22766	31063	-2375
0.5	22460	31090	-21313	50988	-6380	0	-42388	23880	-118335	-76950	-5540	-5575	16188	16188	20825
0.4	-31900	20820	22263	9000	17290	0	-55688	33300	-65814	-48463	-14380	-3950	17813	24500	13350
0.3	-46300	6060	18913	34075	11760	0	-56163	12020	-71665	-25863	-25310	12875	-4063	-10375	20838
0.2	6640	13160	-48988	51250	-17960	7175	-54538	14340	-86684	-9213	-23160	14400	16031	3063	25788
0.1	-19020	9620	-15288	55525	-19500	8012.5	-44638	22920	-95848	25513	-27340	15700	-10438	10063	26663

TABLE 5.3b The Number of Trades Associated with the Performance in Table 5.3a

Box (%)	Crude Oil	Cotton	DAX	EUR	Bund	Euro-dollars	S&P	Gold	Heating Oil	Japanese Yen	NASDAQ	Soy-beans	10-Year Notes	30-Year Bonds	Wheat
2.0	53	41	74	9	0	0	17	53	59	4	17	193	4	16	17
1.9	61	47	76	7	0	0	19	51	59	4	19	203	6	16	21
1.8	71	51	80	11	2	0	17	69	67	8	21	209	4	18	23
1.7	65	51	90	11	2	0	23	64	79	6	27	221	6	16	29
1.6	69	59	106	11	2	0	23	71	81	10	33	225	8	24	31
1.5	83	57	120	17	4	0	23	73	91	12	29	231	14	22	33
1.4	79	63	126	19	6	0	27	82	99	10	41	245	14	24	35
1.3	91	67	134	23	6	0	37	92	111	14	47	251	14	32	41
1.2	109	82	152	23	8	0	39	100	135	14	51	267	18	42	41
1.1	117	88	168	27	6	0	47	112	145	20	67	273	20	44	45
1.0	139	100	188	33	10	0	55	125	163	22	69	297	24	54	55
0.9	141	100	236	37	10	0	61	142	183	32	79	313	32	66	61
0.8	175	116	276	49	14	0	65	154	199	38	97	303	38	76	83
0.7	195	150	314	65	18	0	79	162	229	50	107	307	48	102	97
0.6	225	174	358	101	26	0	117	184	241	60	157	319	52	130	111
0.5	227	198	394	129	38	0	149	192	275	92	187	325	78	169	135
0.4	255	228	418	186	61	0	189	208	273	127	235	325	120	217	167
0.3	273	268	470	231	103	0	241	250	297	179	261	309	172	267	199
0.2	277	290	520	275	187	4	283	278	297	247	287	309	239	289	251
0.1	285	290	546	293	285	8	299	282	297	293	293	307	294	297	285

3-box reversal for a total directional reversal of $30. In Table 5.3a, all negative net profits are shaded in grey, and the best results for each market are outlined.

Results show inconsistency. The more trending markets, the interest rates and the euro, are generally profitable for all box sizes. The noisier markets, primarily the equity index markets, show consistent losing results. Eurodollar interest rates show few trades because a percentage of the price at 99.00 is far too big for a box size. To get the right box size, the series would need to be converted to yields. The same is true of the longer-term rates, even though both 10-year notes and 30-year bonds showed good returns. The grains were also inconsistent, with soybeans showing too many trades (see Table 5.3b) while wheat appears normal. This reflects the volatility of the individual markets.

Even during this 10-year test, the volatility of each market would have changed considerably. Systems that are event driven, such as swings, point-and-figure, and breakouts, are not bothered by low volatility for moderate periods of time. They typically hold the same position until something new happens. However, if at the start of the test the volatility was much lower than at the end, there will be few trades at the beginning and much more at the end. Results will be distorted, with greater importance given to the more volatile periods.

It seems reasonable to conclude that the change in price and, consequently, the change in volatility determine the most practical choice of box size. As an example, if we look at the best choice of box size for a selection of stocks and industrial groups for 2003, we get the results shown in Table 5.4. These are plotted as a scatter diagram in Figure 5.14. They show a clear relationship between price level and box size. We will use this pattern to create a more dynamic point-and-figure chart. Rules for varying box sizes and risks associated with these price and volatility changes are discussed later in the section, "Point-and-Figure with Variable Box Size."

TABLE 5.4   Point-and-Figure Box Sizes—Stocks and Sectors 2003

Market	Box Size	Average Price
AMZN	13	25
AMR	38	20
AOL	6	13.5
GE	25	27
IBM	143	72
INTC	80	17.5
MRK	162	52
XOM	50	34
Aerospace	17.5	1,800
Biotech	15	700
Large banks	8	1,000
Life Insurance	22	1,450
Semiconductors	1.6	300

Point-and-Figure Trading Techniques

The basic point-and-figure trading signals are triggered on new highs and new lows:

Buy when the filled column of *X*s, the current upswing, rises above the previous
column of *X*s by one box.

Sell when the filled column of *O*s, the current downswing, falls below the low of the
previous column of *O*s by one box.

Using the basic rules, you are always in the market, reversing from long to short and
from short to long, unless you only trade stocks from the long side.

There are alternate methods for selecting point-and-figure entry and exit points that
have become popular. Buying or selling on a pullback after an initial point-and-figure sig-
nal is one of the more common system entries because it can limit risk and still maintain
a logical stop-loss point. Of course, there are fewer opportunities to trade when only
small risk is allowed, and there is a proportionately greater chance that the trade will
be stopped out because the entry and exit points are close together. There are three ap-
proaches recommended for entering on limited risk:

1. *Wait for a reversal back to within an acceptable risk, then buy or sell immediately
   with the normal point-and-figure stop.* Figure 5.15 shows various levels of risk in IBM
   with $2.00 boxes. The initial buy signal is at $150, with the simple sell signal for liquida-
   tion at $134, giving a risk of $16 per share. Instead of buying as prices reach new highs,
   wait for a reversal after the buy signal, then buy when the low for the day penetrates
   the box corresponding to your acceptable level of risk. Three possibilities are shown in
   Figure 5.16. Buying into a declining market assumes that the support level (at $134 in
   this example) will hold, preventing the stop-loss level from being reached. To increase
   confidence, the base of the formation should be as broad as possible. The test of a
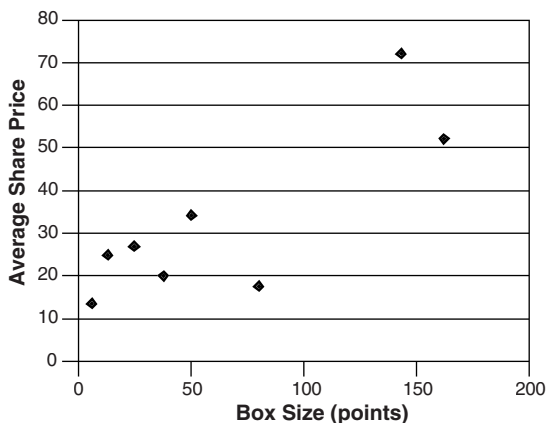
FIGURE 5.15 Point-and-figure box sizes are larger when prices are higher
or volatility is greater.

```
Price  _____
 160
 158   o
 156   o
 154   o                      _____
 152   o  x
 150   o  x  o              | x |   <----- Simple buy signal, standard risk of $16
 148   o  x  o  x             x
 146   o  x  o  x  o  x       <----- Enter on pullback to 146, risk $12
 144   o  x  o  x  o  x
 142   o  x  o  x  o  x       <----- Enter on pullback to 142, risk $8
 140   o  x  o  x  o  x
 138   o     o  x  o  x       <----- Enter on pullback to 138, risk $4
 136         o     o
 134                  [    ]  <----- Stop-loss and reversal at $134
```
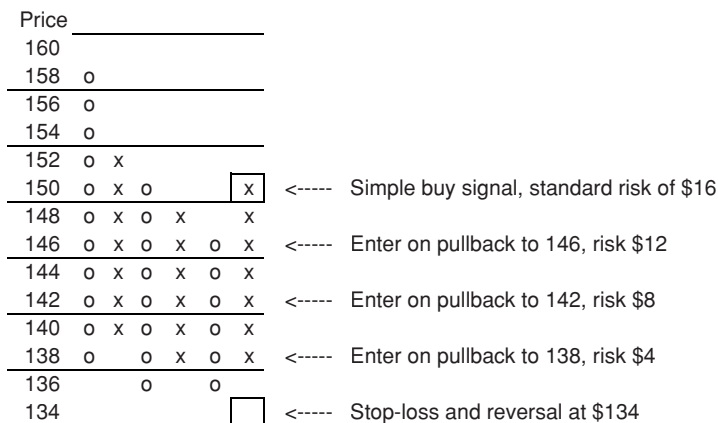
FIGURE 5.16 Entering IBM on a pullback with limited risk.

triple bottom or a spread triple bottom after a buy signal is a more reasonable place to go long than a simple buy after a small reversal in the middle of a move.

It is not advisable to reduce risk by entering on the simple buy signal and placing a stop-loss at the point of the first reversal (3 boxes below the highs). The advantage of waiting for the pullback is that it uses the logical support level as a stop. A stop-loss placed nearby following a breakout has no logical basis and will quickly result in a losing trade.

2. *Enter the market on the second reversal back in the direction of the original signal.* As shown in Figure 5.17, the first reversal following a signal may not reach the target risk level. Price movement does not often accommodate our expectations; therefore, a more flexible rule is needed. One technique is to enter a long position on the second upswing. That is, do not enter a long position on the initial buy signal but wait until a 3-box reversal has caused a downswing. As the downswing continues, place a trailing buy order at the point where the next upswing would begin, at the fourth box above the lowest box of *O*s. If the order is executed, then there is a new position
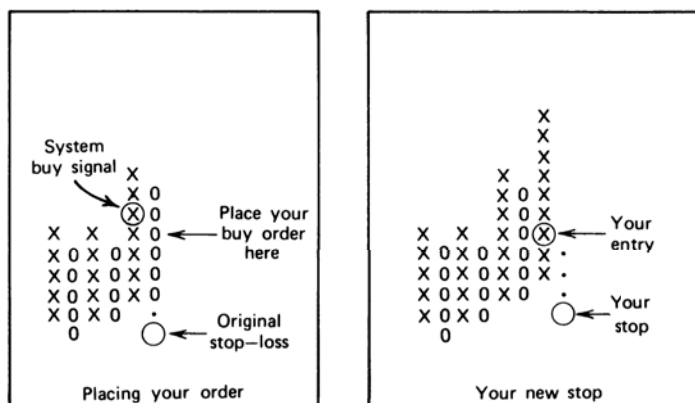
FIGURE 5.17 Entering on a confirmation of a new trend after a pullback.

in the direction of the trend; however, the risk has been limited to the value of four boxes, which is the new trend reversal point. In a wide-ranging, volatile market, this entry may be better than the original point-and-figure buy signal.

This technique is frequently used by traders, who firmly believe that a reversal follows immediately after a breakout, to prevent both high risk and false signals. If the pullback that follows the breakout continues in an adverse direction, penetrates the other support or resistance level, and triggers the original system stop-loss, then no entry occurs, thus saving a substantial whipsaw loss. These traders are essentially looking for a confirmation of direction. However, if prices continue upward, without a pullback, the trade may be missed entirely. You can read more on this topic in the section "Individual Trade Risk" in Chapter 23.

The reversal principle in Step 2 can also be effective for building positions. In bar charting (discussed in Chapter 3), a pullback to a bullish support line or a bearish resistance line is a point for adding to a position with a risk limited to penetration of the major trendline. The equivalent procedure using point-and-figure is to add on each reversal back in the direction of the trend using the newly formed stop-loss point to exit the entire position (as shown in Figure 5.18).

3. *Allowing for irregular patterns.* Price patterns are not always orderly, and the price activity at the time of a trend change can be very indecisive. One basic trading principle demands that the market confirm a new high before buying; the first new high might simply occur during an erratic sideways pattern, or an expanding formation after a period of low volatility. If prices are required to make a new high by more than one box on the next upward thrust, and raise the lowest point of the reversal by another box on the third upward thrust, we are actually demanding that the momentum, or speed of price movement, increase before a position is set.[10] The pattern of higher highs and higher lows is similar to upwards acceleration.
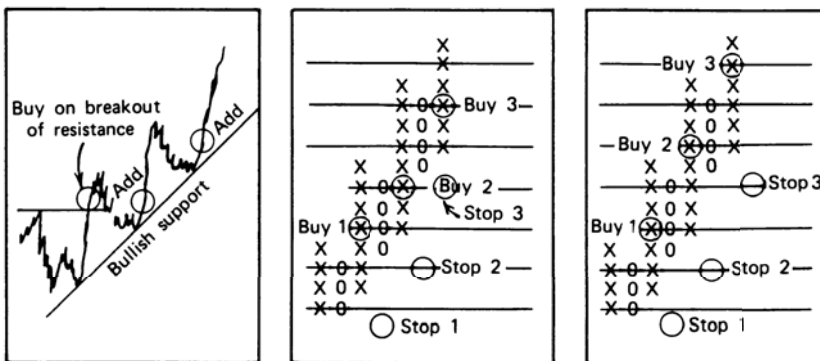
FIGURE 5.18 Three ways to compound positions.

---

[10] Adam Hewison, "The Will Rogers Theory of Point & Figure Trading," *Technical Analysis of Stocks & Commodities* (August 1991).

This technique, which tends to minimize false breakouts, may be modified to increase the confirmation threshold from two to three or four boxes as market volatility increases. If the box size is changed according to volatility, as discussed later in this chapter, confirmation can remain at two boxes.

Point-and-Figure Trading Risk

I go long or short as close as I can to the danger point, and if the danger becomes real I close out and take a small loss.

—Jesse Livermore to Richard Wyckoff[11]

The point-and-figure method is a simple trend-following concept, yet in its normal use it is subject to interpretation using trendlines and geometric formations. Why is that necessary? The answer involves the risk of an individual trade. In the previous section, the risk of a trade was changed by using alternate entry points based on pullbacks; however, there are other choices.

The treatment of the same price move can be seen by looking at both a bar chart and a corresponding point-and-figure chart for the same period (Figure 5.19). The most basic bar chart trading method uses horizontal support and resistance lines to define a trading range; when the resistance line is penetrated, a long position is entered. This is the same concept used in the swing method. A stop-loss is placed below the resistance line in order to close out the trade in the event of a false breakout. An alternate placement of the stop-loss could have been below the support line, allowing the new bull move some latitude to develop.
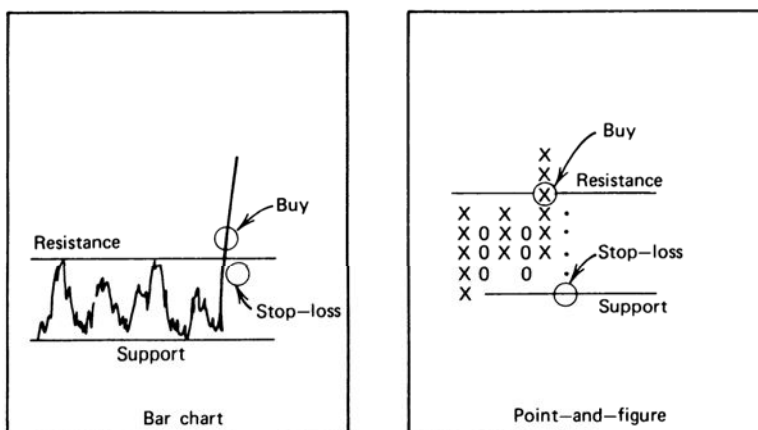
FIGURE 5.19   Placement of point-and-figure stops.

[11]Richard D. Wyckoff, *Stock Market Technique, Number One* (New York: Wyckoff, 1933), 2.

Viewing this chart pattern in retrospect makes the selection of the entry point and the placement of stops or reversal of position seem obvious; however, when trading, the choice of the support and resistance lines is not usually as clear. The time to enter a trade after a breakout is never quite certain, and the position for the stop or reversal depends on the volatility of prices and the risk that you are willing to take. In contrast to the ambiguity of the bar chart, the point-and-figure method defines the support and resistance levels exactly, establishes a place to buy in advance, and designates the position for the stop-loss below the rectangular congestion area, always at the place where a long position would turn to a short sale. The rigidity of the method allows only one place for the stop-loss and fixes the risk as the difference between the support and resistance lines, a total of five boxes in the example shown in Figure 5.19. In the bar chart, the risk might have been held to the equivalent of two boxes using the closer stop-loss, where the trade is exited on a pullback into the range; however, a very small risk often results in being stopped out of the trend prematurely.

Take It and Run!

From time to time, you find yourself the beneficiary of a substantial price move where there is an uncomfortably large profit. It is normal to consider how much of the unrealized profits will be returned before the system finally generates an exit signal. At these times, some traders prefer to take the profits. These are decisions that go beyond the area of technical analysis, although they could be rigorously tested. If the profit currently held in open positions is enough to sustain a life of leisure, a home in the mountains, membership in a country club, and a small investment in a hotel or restaurant to occupy your time, then take it and run! A trading system should not depend on a single, very large profit to work over time. It should have a steady, successful profile. We can therefore reason that any extremely large profit is a windfall and should be taken. Occasionally, a price shock gives you a windfall profit that has nothing to do with a well-designed system or astute trading. It is another opportunity to take profits. If you are correct, prices will reverse after you have gotten out, and there will be an opportunity to reenter the trade at a much better level.

If you cannot sleep nights and the unrealized profits are just enough to satisfy some important obligations, but one or two adverse days would ruin the opportunity, then take the profits and begin again with a small investment. While it is important to follow a system in order to benefit from the long-term expectations, investor "risk preference" preempts the act of blindly following all rules. Sometimes you need to reduce the market exposure, whether the current position is a profit or a loss. If you have some risk latitude, a logical place to capture most of the current open profits and still have a chance of increasing those profits is to use the point-and-figure reversal value for your exit. If the 3-box reversal represents only 10% of the open profits, then you have a significant chance of profiting from a continuation of the price move. The *reversal value*, the box size times the number of boxes in a reversal (usually 3), is meant to indicate a significant contrary move and can be used as an objective indication of a change of direction. One approach to taking profits is shown in Figure 5.20.
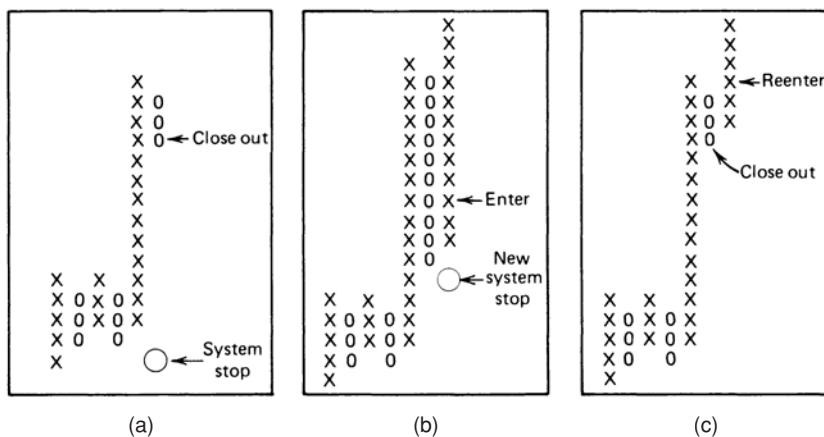
FIGURE 5.20   Cashing in on profits.

In Figure 5.20a, a trailing 3-box reversal value is used for the stop-loss once there is a sustained move of at least 10 boxes. To reenter the move in the same direction, the same technique is used (Figure 5.20b), adding another 4 boxes of profits while keeping the new risk small. In Figure 5.20c, this method lost 4 boxes of the potential profits when the reversal was short lived.

In general, taking small profits in this way does not improve overall profitability because it most often misses the biggest moves. It may, however, reduce the risk of loss at a rate even faster than the reduction in profits, yielding a better reward-to-risk ratio. A new rule should be carefully tested to know its effect on different market conditions before it is used. This topic is covered thoroughly in Chapter 21.

Control of risk can also be accomplished effectively using *volatility stabilization* methods. These continually adjust your position size to maintain a *target volatility*. If the volatility of either a market or your portfolio increases sharply, volatility stabilization will cause positions to be cut in order to bring your risk down to an acceptable level. See Chapter 23 for more about this.

Alternate Treatment of Reversals

Traditional point-and-figure charting favors the continuing trend. On highly volatile days, *broad ranging days* (also called *outside days*), it is possible for both a trend continuation and a 3-box reversal to occur. Point-and-figure rules require that the trend continuation be recorded and the reversal ignored. Figure 5.21 shows a comparison of the two choices. In the example, prices are in an uptrend when a new 1-box high and a 3-box reversal both occur on Day 6, as seen in Figure 5.21a. In Figure 5.21b, the traditional approach is taken, resulting in a continuous upward trend with a stop-loss at 7.90. Taking the reversal first as an alternate rule, Figure 5.21c shows the same trend with a stop-loss now at 8.05, 15 points closer.

Plotting the reversal first usually works to the benefit of the trader; both the stop-loss and change of trend will occur sooner. Subsequent computer testing proved this to be

Day	High	Low
1	$8.02	$7.90
2	$8.11	$7.95
3	$8.13	$7.92
4	$8.18	$7.96
5	$8.25	$8.10
6	$8.32	$8.08
7	$8.40	$8.25

```
       $8.40             X                        X
       $8.35             X                        X
       $8.30             X                        X
       $8.25             X                   X    X
       $8.20             X                   X O  X
       $8.15             X                   X O  X
       $8.10  X          X           X       X O
       $8.05  X O        X           X O     X
       $8.00  X O        X           X O     X
       $7.95  X O                    X O
       $7.90  X                      X

            (a)        (b)                  (c)
```

FIGURE 5.21 Alternate methods of plotting point-and-figure reversals. (a) Sample prices for plotting. (b) Traditional method. (c) Alternate rule taken on Day 6.

true. This alternative does not work when the reversal value is small and broad ranging days (that trigger the optional reversals) occur frequently.

Selecting Trades

Not all trades are profitable in any trading system. Some analysts prefer point-and-figure charts because both the profit objective and the risk can be identified at the time of entry. The profit objective can be calculated using the vertical or horizontal count, and the risk is the size of the price reversal needed to cause the opposite signal. Trades are then taken only if the return to risk ratio is greater than 2.0.

As with other trending systems, 45° trendlines can be drawn to identify the current dominant trend. Trades may be taken only in the direction of that trend. In a bull market, new short signals are ignored until the box is filled that penetrates the upwards bullish trendline. Then the bias switches to the short side.

Price Objectives

Point-and-figure charting has two unique methods for calculating price objectives: *horizontal counts* and *vertical counts*. These techniques do not eliminate the use of the standard bar charting objectives, such as support and resistance levels, which apply here as well.

The Horizontal Count

The time that prices spend in a consolidation area is considered directly related to the size of the subsequent price move. One technique for calculating price objectives is to measure the width of the consolidation (the number of columns on a bar chart) and project the same measurement up or down as the target of the move. The point-and-figure horizontal count method is a more exact approach to the same idea.

The upside price objective is calculated as

$$H_U = P_L + (W \times R)$$

where H_U = the upside horizontal count price objective
            P_L = the price of the lowest box of the base formation
            W = the width of the bottom formation (number of columns)
            R = the reversal value (number of boxes times the value of one box)

To use this formula, the *base* (width of the bottom or top formation) needs to be identified. Count the number of columns, W, not including the breakout column and multiply that width by the value of a minimum reversal, R; then add that result to the bottom point of the base to get the upper price objective. The base can always be identified after the breakout has occurred. For example, Figure 5.22 shows the March 74 contract of London Cocoa (£4 box) forming a very long but clear base. The reversal value is £12, and the width of the base is 19 columns (not counting the last column, which included the breakout). Added to the lowest point of the base (£570) this gives an objective of £798, reached on the left shoulder of the topping formation. Another alternative is the wider base, marked as $W_2 = 25$. Using this selection results in a price objective of £870, by adding $25 \times £12 = £300$ to £570, the lowest point of the base.

The downside objective is calculated in the same manner as the upside objective:

$$H_D = P_H - (W \times R)$$

where H_D = the downside horizontal count price objective
            P_H = the price of the highest box of the top formation
            W = the width of the top formation (number of columns)
            R = the reversal value

Some examples are given for downside objectives in the same Cocoa diagram (Figure 5.22). A small correction top could be isolated at the £720 level and two possible top widths, W_3 and W_5, could be chosen. The broader top, W_3, has a width of nine and a downside objective of £632. The shorter top, W_5, has a width of five and a downside objective of £680. Although the closer objective, calculated from W_5, is easy to reach, the farther one is reasonable because it coincides with a strong intermediate support level at about £640.

The very top formation, W_4, was small and only produced a nearby price objective similar to the first downside example; there would be no indication that prices were ready for a major reversal. The top also forms a clear head-and-shoulders pattern, which could be used in the same manner as in bar charting to find an objective. Using that technique, the distance from the top of the head to the point on the neckline directly below is 20 boxes; then downside price objective is 20 boxes below the point where the neckline was penetrated by the breakout of the right shoulder, at £776, giving £696 as the objective.
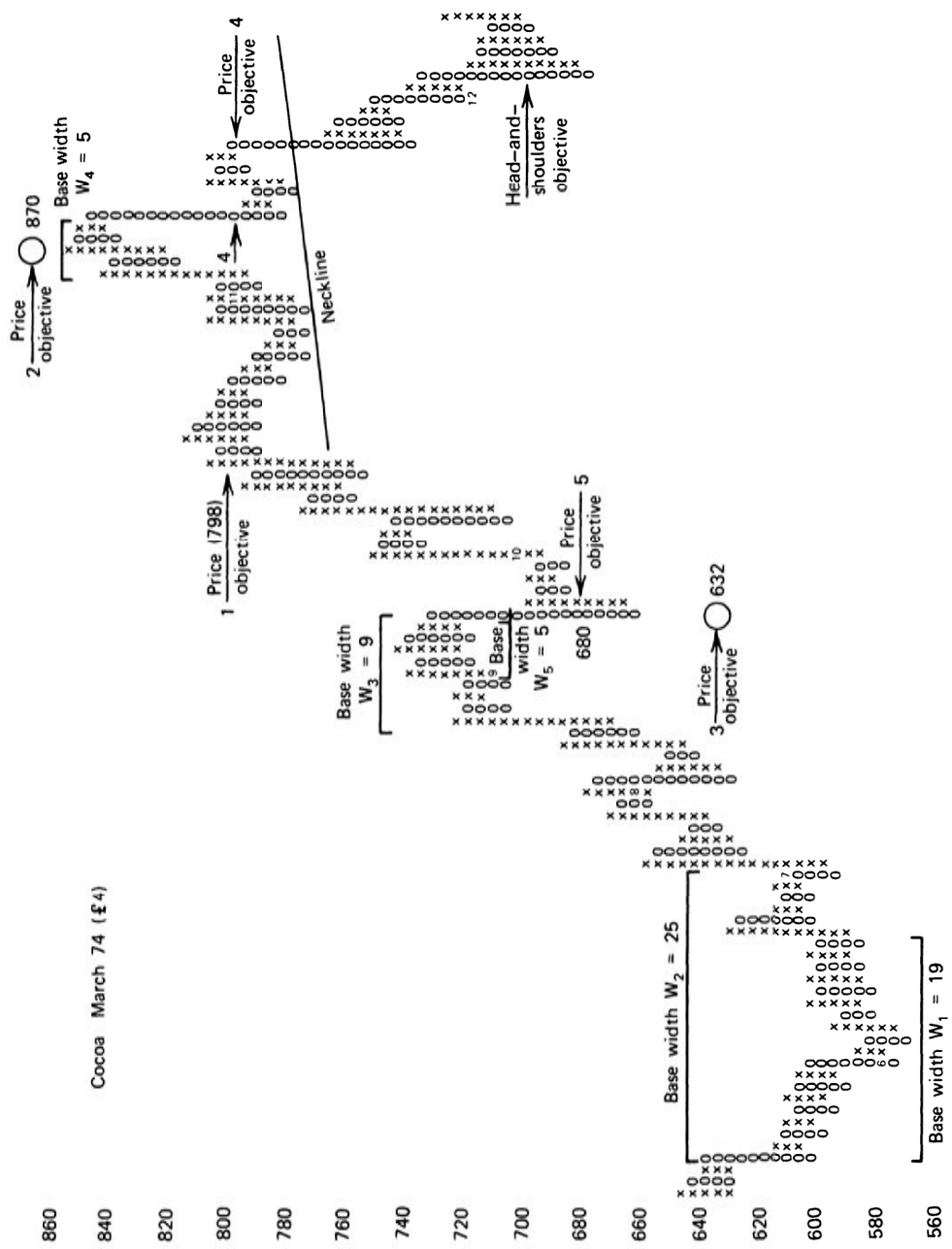
FIGURE 5.22 Horizontal count price objectives.

The horizontal count can also be applied to a breakout from a triangular formation, similar to the one on the very far right in Figure 5.22 (marked "Head-and-shoulders objective"). The width of the formation is the widest point in the center of the triangle, and the upwards objective is also measured from the center, rather than from the bottom of the triangle.

The Vertical Count

The vertical count is a simpler and more definitive calculation than the horizontal count. As with the horizontal count, there is adequate time to identify the formations and establish a price objective before it is reached. The vertical count is a measure of volatility (the amount of rebound from a top or bottom) and can be used to determine the size of a retracement after a major price move. To calculate the upside vertical count price objective, locate the first reversal column after a bottom formation. To do this, a bottom must be established with one or more tests, or a major resistance line must be broken to indicate that a bottom is in place. The vertical count price objective is then calculated:

$$V_{up} = Lowest\ box + (First\ reversal\ boxes \times Minimum\ reversal\ boxes)$$

The downside vertical count price objective is just the opposite:

$$V_{down} = Highest\ box - (First\ reversal\ boxes \times Minimum\ reversal\ boxes)$$

where   *First reversal boxes* = the number of boxes in the first reversal
        Minimum reversal boxes = the number of boxes needed for a chart reversal

Examples illustrating the vertical count are easy to find. In the QQQ chart (Figure 5.23), the NASDAQ low is clearly in early October 2002, followed by an upwards reversal of 13 boxes. Each box is $0.25, giving a total reversal of $3.25. Multiply the reversal amount by 3, the number of boxes in a reversal, and add that value, 9.75, to the low of $20.00 to get the target of $29.75. That value was reached during May 2003.

A secondary low in the QQQ chart occurs in February 2003 at $23.50. The first reversal that follows is seven boxes, or $1.75. Multiple 1.75 by 3 and add the result to the low to get the target of $28.75. The two objectives confirm each other. As a simple measurement tool based on recognizing key highs or lows, the vertical count relies on volatility to determine the extent of the move that follows. It can be quite accurate at times; otherwise it is likely to understate the expected price move.

In Figure 5.24, Intel is used to show frequent possibilities for the point-and-figure vertical count. As with QQQ, the chart uses a box size of 25 points ($0.25) and a 3-box reversal. The low in August 2002 at $19.00 gives a target of $27.25 reached at the end of October. The high in early January 2003 points to a very nice interim low in April, but falls far short of the actual low. Two other midrange price reversals could be used to find additional downside targets, each of which is successful but also fall short of predicting the true low of the stock. The vertical count can be very reliable but should not be expected to be a magic wand.
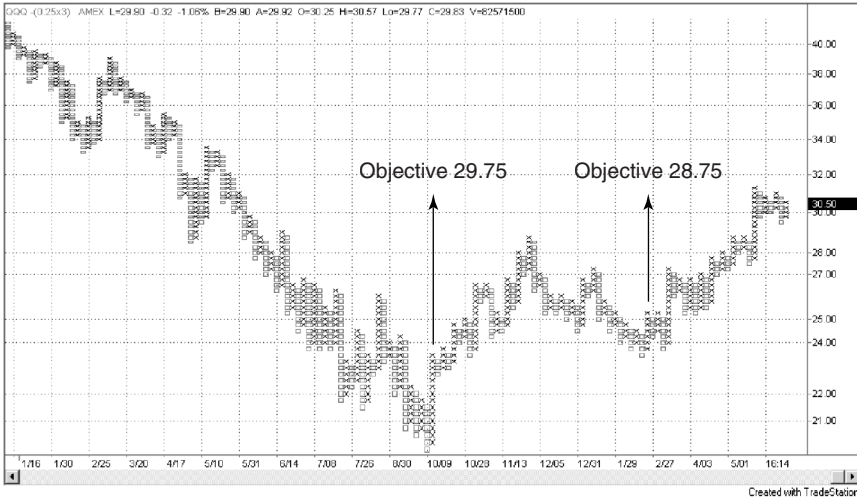
FIGURE 5.23 Point-and-figure vertical count for QQQ. The major low in October 2002 had a reversal of 13 boxes, each with a value of 0.25, for a total of 3.25. Multiplying by 3 and adding the result, 9.75, to the lowest value $20.00, gives the price objective of 29.75. The second bottom in February 2003 shows a reversal of 7 boxes and gives a target of 28.75.
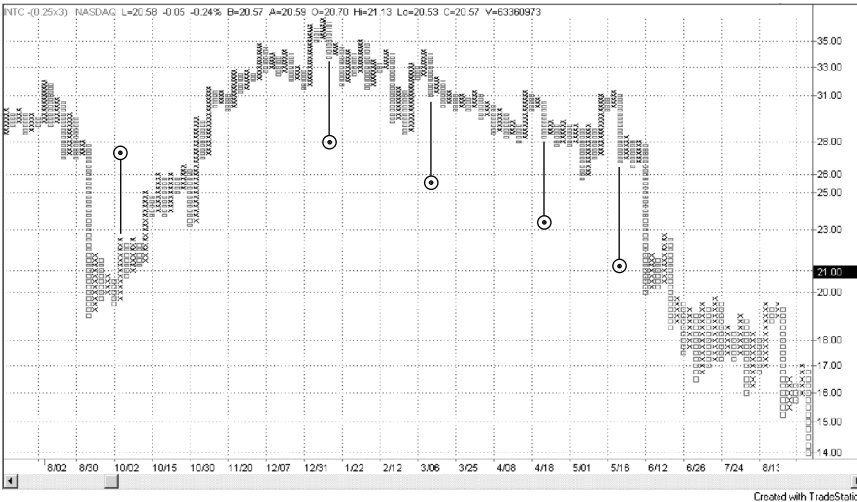
FIGURE 5.24 Intel vertical count examples. This chart of Intel had many possibilities for the point-and-figure vertical count, all of which understate the actual price move. The vertical count is a good tool for setting initial objectives, but it is not a magic wand.
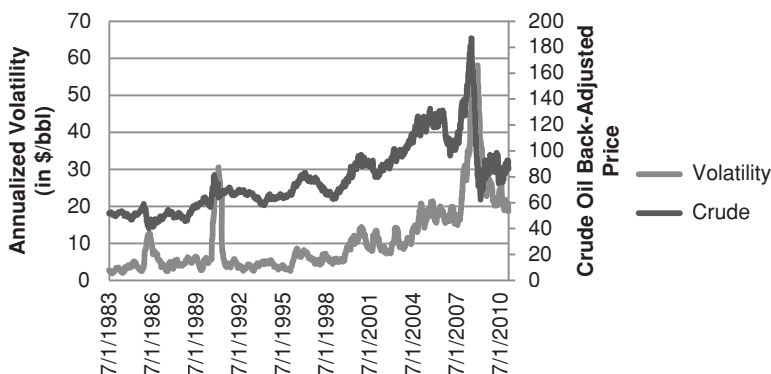
FIGURE 5.25 Annualized volatility of crude oil in $/bbl using back-adjusted futures.

Point-and-Figure with Variable Box Sizes

In the section "Point-and-Figure Box Size," we looked at the results of various fixed point box sizes for a selection of markets. An optimization showed the net profits from boxes varying from 0.10% to 2.0%, all using a 3-box reversal (see Table 5.3). The inconsistency of results shows that each market must be treated uniquely, or that a more general approach to volatility-adjusting the box sizes is needed. Volatility will be discussed in detail in Chapter 20; however, some simple principles can be used here. Two possible solutions that would change the box size daily (or periodically) are based on

1. A fixed percentage of the current price
2. A fixed multiple of the average true range

These two methods will be applied to three futures markets, the euro, S&P, and gold, for the period from 2000 through 2010. That interval had every combination of price movement, strong trends, both down (after the tech bubble) and up, as well as the most severe economic crisis in many years. The price of gold more than tripled as well. Each showed very different results in Table 5.3.

To further support the need for a more flexible approach, consider crude oil, shown in Figure 5.25. The right scale shows the price of continuous futures (the cash price peaked near $150/bbl) and the left scale the annualized volatility in percent. With a price range of $30 to $150 during the past 10 years, and a volatility of less than 10% to near 60%, no one choice of box size will work. If you find a successful box size for 2007, it would be too large to generate any orders before 2000. If you used the best box size for the 1990s, it would be changing signals every day in 2007.

Boxes Based on a Fixed Percentage of Price

Volatility is directly related to price; therefore, we can use a percentage of the current price to adjust the box size. For gold at $500, a 0.5% box size gives a $2.50 box and a $7.50

TABLE 5.5 Results of Using Point-and-Figure with Variable Box Size Based on a Fixed Percentage Futures Data from 2000 through 2010

Box (%)	Net Profits or Losses			Number of Trades		
	EUR	**S&P**	**Gold**	**EUR**	**S&P**	**Gold**
2.0	0	19,538	13,770	0	21	6
1.9	0	30,650	51,160	0	19	8
1.8	0	−28,875	41,260	0	23	8
1.7	0	−29,038	11,4440	0	28	10
1.6	0	7,150	6,100	0	37	16
1.5	0	11,625	55,120	0	45	22
1.4	0	−14,300	89,540	0	45	24
1.3	54,900	56,050	−52,000	1	57	30
1.2	57,988	85,450	9,180	3	71	31
1.1	49,250	54,800	−33,800	4	77	42
1.0	71,013	58,875	45,520	7	101	62
0.9	15,575	81,525	−57,740	9	129	78
0.8	−100,200	100,200	−54,120	17	149	102
0.7	−73,000	85,400	−46,240	29	185	130
0.6	−30,288	111,500	−61,520	61	245	158
0.5	110,863	93,275	−87,640	110	305	230
0.4	−47,438	115,800	−13,8550	196	385	292
0.3	−55,750	89,075	−13,4300	317	449	380
0.2	46,475	106,650	−13,5380	461	553	498
0.1	46,225	111,100	−12,9900	553	593	556

reversal criteria. At $1500, the box size is $7.50 and the reversal $22.50. With gold highly volatile, and daily ranges greater than $22.50, that seems a bit sensitive, so a 1% box size might be better. Prices may also have been represented on a log scale.

When the box size is varied each day, and the point-and-figure chart is essentially re-drawn, rising price will tend to hold the position longer because the reversal criteria get larger. On the other side, falling price will shorten the reversal and may generate a new signal where there was none yesterday.

Results in Table 5.5 are very different from those in Table 5.3. Where the euro was profitable everywhere in the original test, it has no trades for higher percentages and is profitable in only half of the remaining trades. The S&P, however, has the opposite result. Instead of erratic, mostly losing results, it is highly profitable for box sizes less than 1.4%. Gold has both better and worse results. Overall, this method changes the shape of the results but seems inconsistent. Results were generated using *TSM Point and Figure Variable Box*, available on the Companion Website.

Boxes Based on a Multiple of the Average True Range

A clear problem with the percentage method in the previous section is that each market has a different level of volatility and that volatility changes. A percentage of the closing

TABLE 5.6   Results of Using Point-and-Figure with Variable Box Size Based on a Multiple of the Average True Range (Factor) Futures Data from 2000 through 2010

	Net Profits or Losses			Number of Trades		
Factor	EUR	S&P	Gold	EUR	S&P	Gold
3.0	42,900	−1,800	85,590	5	9	8
2.8	24,913	−2,925	73,070	7	9	10
2.6	38,888	−9,563	76,410	7	11	12
2.4	38,788	3,625	71,030	7	9	8
2.2	−9,313	1,088	23,350	11	11	13
2.0	−77,363	−88	18,740	11	13	14
1.8	−17,538	−5,013	58,440	13	15	18
1.6	−23,688	4,763	33,980	13	15	22
1.4	12,538	5,713	24,070	21	17	24
1.2	31,088	21,150	19,30	31	21	26
1.0	16,963	39,538	−490	33	27	44
0.8	46,463	25,463	31,940	39	41	50
0.6	30,838	14,113	53,170	75	61	68
0.4	55,513	−18,163	14,320	121	127	126
0.2	18,538	−4,588	30,270	253	203	223

price is not sufficient to capture these variations. One measure that has proved valuable is the average true range. In this strategy, the average true range over the past 50 days was multiplied by a factor, shown in the first column of Table 5.6. While the factor does not change, the volatility will vary considerably causing the box sizes to vary.

Results can be seen as much more uniform in both net profits and the number of trades. Rather than no trades in the euro, the three markets all show about the same pattern of trades based on the factor size. Profits also cluster. That is, the losing results are mostly adjacent to one another. All markets, on average, did very well. While this is intended to be an example of the way a strategy can adapt to changing volatility, it seems to show that there are some techniques that will succeed. These results were generated using *TSM Point & Figure ATR Box*, available on the Companion Website.

Finding the True Price-Volatility Relationship

One other method must be discussed. The fixed percentage and ATR approaches were convenient and easy to program. But perhaps the best results will come from finding the true price-volatility relationship. As an example, we calculated the rolling 20-day annualized volatility of crude oil using futures from 1983 to 2010. Figure 5.26 shows that relationship as a scatter diagram. While there are sets of points near the top of the chart, the underlying relationship is very clear, clustering at about 7% when crude was $20 and reaching about 50% when crude was $140. This cannot be found using a linear regression because the higher outliers will cause the regression line to have a steeper ascending angle.
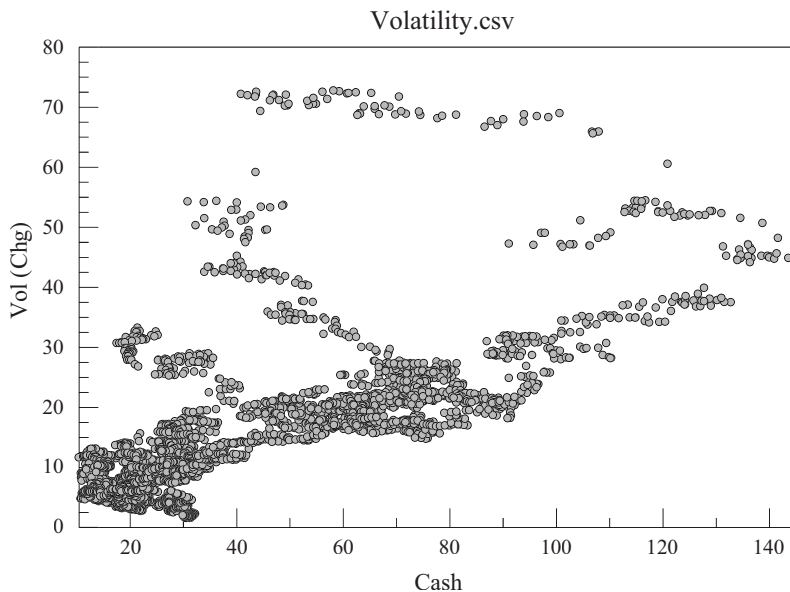
FIGURE 5.26 Volatility relationship for crude oil, 1983–2010 (using PSIplot)

Stock Dividends and Splits

When using point-and-figure charts for stocks, an adjustment must be made whenever a stock dividend is issued or the stock splits, because the chart represents the price of one share. Splits and dividends result in stock multiplying factors:[12]

Activity	Stock Multiplication Factor
10% stock dividend	1.1
30% stock dividend	1.3
2 for 1 stock split	2.0
3 for 2 stock split	1.5

These multiplication factors can be used to correct the box size of a percentage or logarithmic point-and-figure chart by dividing all the boxes by the multiplication factor; therefore, the new box sizes represent the value of one share.

Recent Applications of Point-and-Figure

Not much new has happened to point-and-figure during the past 100 years; however, there has been renewed interest in using it. Two new books, *Power Investing with Sector Funds* by Peter Madlem and *Point & Figure Charting*, second edition, by Thomas Dorsey, show more recent examples of how this technique applies to stocks and sector

---

[12]William G. S. Brown, "Logarithmic Point & Figure Charting," *Technical Analysis of Stocks & Commodities* (July 1995).

indexes. There are also a number of websites with instructions and examples, most combined with advertising.

THE *N*-DAY BREAKOUT

Close on the heels of the swing and point-and-figure methods is the rolling breakout, or *N*-day breakout. In the theoretical sense, it is not entirely an event-driven method because it is affected by time, but its best feature is that it buys on new highs and sells on new lows, very much the same as event-driven techniques. The *N*-day breakout has become one of the most popular trend-following techniques. Its rules are simply:

Buy when today's high moves above the high of the past *N* days.
Sell when today's low moves below the low of the past *N* days.

Using these basic rules you would always be holding a long or short position, reversing direction whenever there is a new signal. A slightly more conservative set of rules is:

Buy when today's close is above the high of the past *N* days.
Sell when today's close is below the low of the past *N* days.

These rules offer a confirmation of the new direction at the cost of setting the new position later than the first set of rules. The intention of the *N*-day breakout is to react immediately to an event that drives prices higher than they have been recently. Figure 5.27 shows a buy and sell breakout in Merck based on a 5-day calculation period. The buy signal occurs in early January 2002 when prices turn up after a decline under $54. The third day up after the double bottom makes a new high above the previous five days. Although there is a wide-ranging day four days later, it fails to make a 5-day low, and the buy signal remains in effect until after prices peak on January 20. Four days later a new low generates a sell signal that is held for the rest of the chart. Two attempts to rally only succeed in making highs above the prior three days, leaving the short sale in place.

This remarkably simple method can be very effective as a trend system. However, the choice of the calculation period, *N*, can change the nature of the technique from one of nearly event-driven to time-driven, as well as from slow-trading to fast-trading. When *N* is less than 5 days, it is difficult to view this method as event-driven; the highs and lows of the few days used are so close together that prices must continue to post new highs in order to continue an uptrend. It is not being triggered by an event but by the continuation of the trend direction.

The performance profile of the *N*-day breakout is very similar to both the swing and point-and-figure methods. Its primary characteristic is that using a large value of *N* causes much greater risk. Risk is defined as the difference between the entry price and the point

MRK TXT-Daily 05/30/2003 C=55.58 +.82 +1.50% O=.00 H=55.80 L=54.90 V=6369600
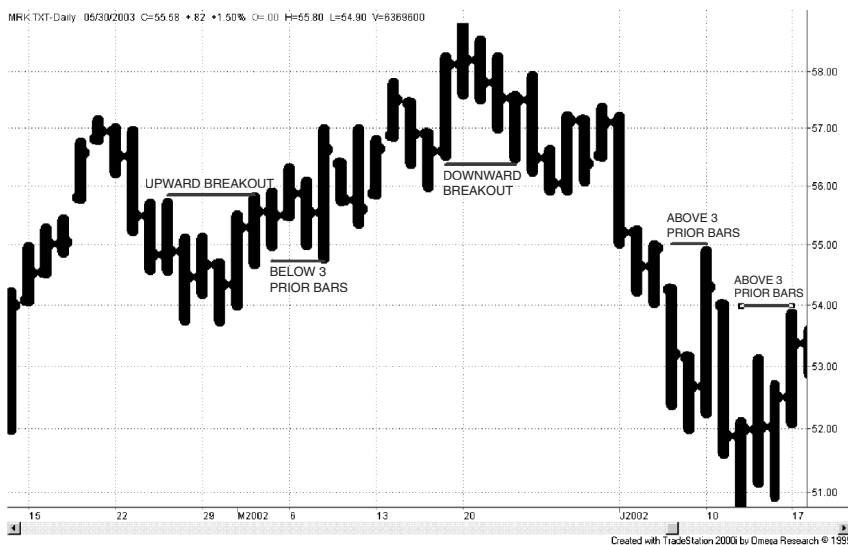
FIGURE 5.27 *N*-Day breakout applied to Merck, using a breakout period of five days.

at which an *N*-day high or low would reverse the signal. This is the primary feature that distinguishes event-driven techniques from time-driven methods, such as a moving average.

A "Comparison of Major Trend Systems," including the *N*-day breakout, can be found in Chapter 8 after a discussion of time-based trends. These characteristics are important to the selection of a trending system, and they vary considerably between event-driven and time-based methods.

Donchian's 4-Week Rule

In 1971, Playboy's *Investment Guide* reviewed Donchian's 4-Week Rule as "childishly simple . . . was recently discovered to rank premiere among a dozen widely followed mechanical techniques." Even now, simple trend-following systems continue to have success. In Donchian's method, the following rules were applied:

1. *Go long* (and cover shorts) when the current price exceeds the highs of the previous four full calendar weeks.
2. *Sell short* (and liquidate longs) when the current price falls below the lows of the previous four full calendar weeks.
3. When trading futures, roll forward if necessary into the next contract on the last day of the month preceding expiration.

In 1970, *The Traders Note Book* rated the 4-Week Rule as the best of the popular systems of the day. Based on 16 years of history, the best performers were December wheat,

June cattle, May copper, August bellies, January soybean oil, and May potatoes. In 1974 there were no financial futures markets.

The system satisfies the basic concepts of trading with the trend, limiting losses, and following well-defined rules. It bears a great resemblance to the principle of Keltner's *Minor-Trend Rule*, modified to avoid trading too often. Over the years this technique, classified as a slow breakout system, has shown the greatest longevity. The system is so simple that the only comments about it must also be simple: Can a 4-Week Rule work for all markets? If price volatility in one market increases dramatically, the high and low for a 4-week period could present enormous risk, while at the same time lower prices could cause extremely narrow ranges in another market. That characteristic is both a key to its success and a problem for risk control. The performance of the more volatile markets, whether good or bad, will overwhelm the quiet market.

In the earlier years of trading, most programs that used futures chose to trade an equal number of contracts. Risk was not the focus, only returns. It is now common to adjust the position sizes, either on entry or dynamically, to maintain a relatively constant exposure to risk, which goes a long way toward correcting the biggest problem in event-driven strategies.

Modifying the *N*-Day Rule

Strategy development platforms, as well as spreadsheets, allow us to take many simple ideas and examine them in great detail, sometimes to excess. We have the ability to backtest the *N*-day breakout to find the best choice of *N* over the test period. We could also find the *N*-week breakout that works best. Carrying the process further, we come to the *n*-minute and *n*-tick rules. And, there may be a benefit for continuing to reduce the time interval. If the 4-Week Rule works but the equity drawdown is too large, shortening the time period could reduce the risk. It may not change the return-to-risk ratio of the method, but it will certainly reduce the absolute size of the profits and losses because trades are held for a much shorter time.

In the *N*-day breakout, the determination of *N* is critical to the success of this system. The most obvious approach to finding *N* is by backtesting a broad range of values (as is shown in the next section). It has also been suggested that *N* could be based on the relationship of normal volatility to current volatility,[13]

$$N_t = N_I \times \frac{V_n}{V_c}$$

where N_t = the number of days used for today's calculation
N_I = the initial number of days used for "normal" markets
V_n = the normal volatility measured over historical data

---

[13] Andrew D. Seidel and Philip M. Ginsberg, *Commodities Trading* (Englewood Cliffs, NJ: Prentice-Hall, 1983).

V_c = the current volatility measured over a fixed period shorter than the period used to define normal volatility, V_n. Typically, this is less than ¼ of the longer period.

As the current volatility increases, the number of days used in today's calculation decreases. This may also be classified as an *adaptive technique*.

Testing the *N*-Day Rule

Using a development platform, it is a simple matter to test the *N*-day breakout on any stock or futures market. In the following example, a selection of popular futures markets was chosen and the data separated into two equal intervals, 1991–2000 and 2001–2010, then combined into 1991–2010. The advantage of using a development platform is that it has an optimization feature that allows the input of a range of calculation periods and outputs the test results for each one so that they can be compared easily. Although testing methods will be discussed in Chapter 21, this process is needed to understand the *N*-day breakout and to introduce concepts that will be used throughout this book.

Using calculation periods from 10 days to 100 days, in increments of 5-days, the *N*-day method was applied to crude oil with the results shown in Figures 5.28a, b, and c. The data used was back-adjusted futures with no costs applied. The first point to note is that, in all cases, there was only one test that resulted in a net loss. The second is that the pattern of success is very different for the first and second 10-year periods, although the combined periods show a tendency to improve as the calculation periods increase.

If we were to select the best choice of calculation period simply by the highest net profit, then the first period would be 75 days, the second 15 days, and the combined period 100 days, showing that the selection process may be more complicated than it first appears. This also will be covered in Chapter 21. Although the *N*-day breakout is a powerful trending tool, its performance for a sample of stocks and futures markets shows that individual market patterns cause the best calculation periods to be different. Table 5.7 is a summary of a varied selection of futures and stocks, showing that the best selection jumps from very short periods to very long, and the robustness of the tests (judged by the percentage of profitable tests) also varies considerably. Finding a robust solution to this testing process will also be discussed in Chapter 21.

Weekly Breakouts

Weekly breakouts, a limited case of the *N*-Day Rule, were tested by Hochheimer and summarized in his report as the *Weekly Price Channel*. Weekly breakouts produce a slow trading system that depends on major moves for profits. About 75% of those futures markets tested in this 1982 study[14] showed higher risk (some very high) than the crossover systems tested earlier (also see Chapter 21).

---

[14]Frank L. Hochheimer, *Computerized Trading Techniques 1982* (New York: Merrill Lynch Commodities, 1982).
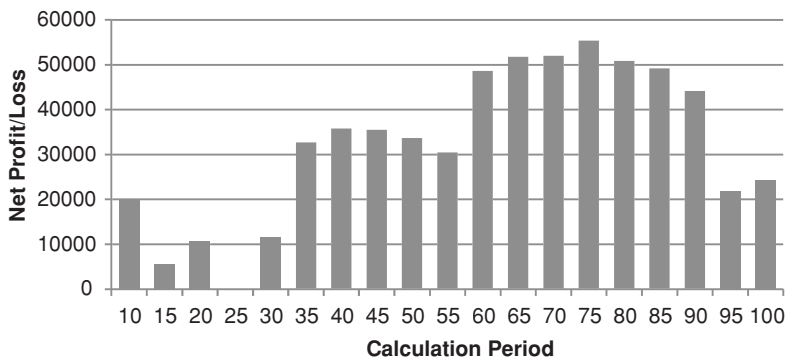
FIGURE 5.28a Crude oil *N*-day breakout optimization, 1991–2000.
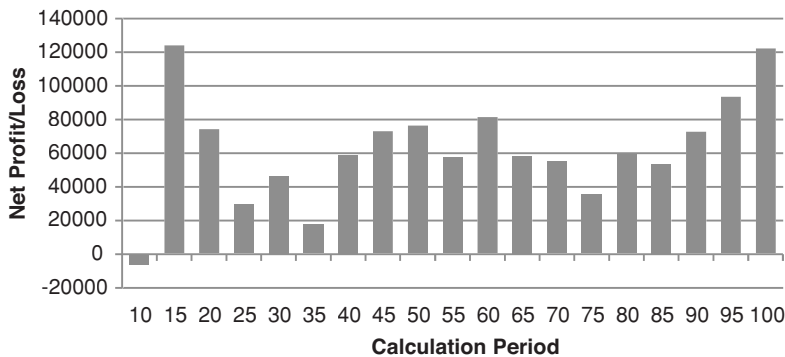
FIGURE 5.28b Crude oil *N*-day breakout optimization, 2001–2010.
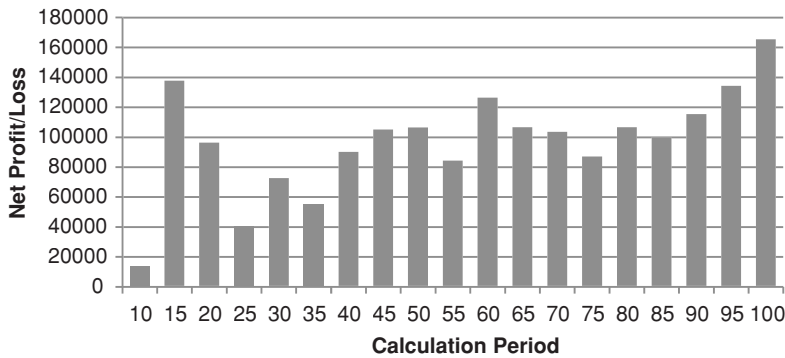
FIGURE 5.28c Crude oil *N*-day breakout optimization, 1991–2010.

TABLE 5.7 Best Results Using *N*-Day Breakout Optimization 2001–2010

	Best Period	Net PL	No. of Trades	% Profitable Tests	% Profitable Trades
Futures					
Crude oil	15	$123,890	86	95%	38%
Euro currency	20	$90,662	59	100%	48%
Gold	90	$15,370	10	21%	31%
S&P Index	65	$44,225	20	79%	39%
US 30-year bonds	20	$27,968	67	63%	40%
Stocks					
Amazon	70	$31.28	20	26%	33%
Bank of America	65	$85.27	16	100%	38%
Caterpillar	65	$104.40	20	79%	42%
General Electric	10	$116.69	146	47%	35%
Hewlett-Packard	95	$39.30	9	95%	39%
Merck	95	$71.52	10	89%	45%
Microsoft	15	$39.11	86	47%	41%

The original purpose for the Weekly Rule was to look at prices only on Friday. The close on Friday is considered important because it is the evening up at the end of the week, in the same sense as the daily close is considered the most important price of the day because all accounts are settled at that price. Many traders have the opinion that holding a position over a weekend is the only thing worse than holding it overnight. This evening-up process is expected to prevent false signals that may occur midday or mid-week during periods of low liquidity. Weekly data are also smoother than daily data, and often show a clearer trend.

Typical trading rules for a Weekly Price Channel system use the more conservative choice, confirming the direction based on closing prices:

- *Buy* (and close out shorts) if the closing price on Friday exceeds the highest closing price of the past *N* weeks.
- *Sell short* (and close out longs) if the close on Friday is below the lowest closing price of the past *N* weeks.

The main drawback of this model is that the risk can be very high. The initial risk of a new long position is the difference between the highest and lowest closing prices of the past *N* weeks. In addition, even if penetrated, the position is not liquidated until the close of Friday. This could be a very risky proposition; however, that risk is offset by the smoothing effect of only acting once each week. It may be that accepting higher risk is better than being subjected to more frequent false signals. When a series of individual trade losses is viewed as a single, larger loss, the risk of weekly positions may not seem quite so large. The risk relative to your investment is also smaller for stocks compared to futures. Leverage in futures, which can be as high as 20 to 1, magnifies the risk of each trade.

Alternate rules were tested by Hochheimer in his study. Buy and sell signals were taken at the time that the intraday new high or new low occurred. No comparison was available to determine whether the risk was greater or less than the conventional approach. Hochheimer also tested the Weekly Rule with the week ending on days other than Friday. It was not apparent that any one day was better; however, the week ending Friday seems to be supported by liquidation before the weekend.

Traders will find the basic breakout method is one of the important trending techniques, and it is used throughout this book as a benchmark study. In Chapter 8 it is included as one of the basic trending methods in "Comparison of Major Trend Systems," and shows 10 years of results for the Eurodollar. A spreadsheet is also available on the Companion Website, *TSM N-Day breakout* example using BA, that allows testing of this method.

Avoiding Problems Programming the Weekly Breakout

The Weekly Rule is often thought of as having signals only on Friday; however, when programming this method, it is important to remember that a number of weeks end on Thursday due to holidays. Identifying the last day of the week in advance is a problem for a computer program. Although we can find the day of the week by simply using the function *DayOfWeek* in TradeStation, or *WEEKDAY* in Excel, we have already missed the day if the information returned jumps from Thursday to Monday and we discover the gap on Monday.

To avoid this problem, many charting systems provide weekly data on request, rather than daily. The automatic built-in conversion program will correctly change days to weeks. By looking at the high, low, and close of the weekly bar, you can decide whether a signal has occurred. Execution will be on the last close of the week, regardless of the day on which it occurs. It is best not to try to use daily data to test a weekly strategy.

Dynamic Breakout System

While the *N*-day breakout adapts to price moves and volatility, placing orders for entry and exit one day ahead may improve timing. In his *Dynamic Breakout System* (*DBS*),[15] Stridsman anticipates the entry and exit points based on a factor of the standard deviation of recent prices, then places stop orders for the next trading day. He also proposes that the study of when the exit stop is hit, that is, the number of days from the position entry, will show that trades are separated into two groups, those that get stopped out very quickly and the others that hold the trade for a much longer time. From this information, he expects traders to position their stops better. A program that generates this information, *TSM Dynamic Breakout System*, is available on the Companion Website.

The *N*-Day Breakout Applied to Stocks

Most literature on breakout systems applies this technique to futures markets. The primary differences between futures and stocks are that futures markets are generally less

---

[15]Thomas Stridsman, "Revelation Trading," *Futures* (February 1998).

correlated and have higher leverage. The normal leverage intrinsic to futures trading will cause a small percentage move to yield a relatively high return. In addition, the cost of trading futures is almost negligible compared to the face value of the contract. For example, one S&P *e*mini contract costs less than $5 per trade. At an S&P price of 1,400 and a value of $50 per big point, one contract has a value of $70,000, making the commission cost about 0.7 basis points (0.0070). While you can trade 1,000 shares of a stock for a flat fee of $9, most stocks average less than $30 per share, about 3.0 basis points. Because the commission is fixed regardless of the number of shares traded, a smaller order will cost more. To decide the success of trading either futures or stocks, it is necessary to show the returns in terms of dollars per contract or cents per share. Only then can you decide if commissions and slippage can be covered comfortably and still net a profit.

Donchian's 40/20 Channel Breakout

Before we applaud newer developers for innovation, we must give the originators credit. Richard Donchian created what now seem to be very simple trading methods, but in the 1960s while working at Hayden Stone, he implemented a number of technical systems that were groundbreaking. They included moving average and breakout systems. At the time, any systematic trend-following method was state-of-the-art and profitable.

The 40/20 Channel Breakout was the earliest *N*-day breakout that is recorded. The longer period was used for entries and the shorter for exits, very much like the Turtles method that follows.

The Turtles

During the mid-1980s the group known as the *Turtles*, founded by Richard Dennis and Bill Eckhardt, was the biggest trading sensation, much like Monroe Trout in the later 1980s and Renaissance Asset Management in the 1990s. They all maintained a high degree of secrecy with regard to their trading methods. Years after the group disbanded the rules and operations of the Turtles became public through publications of Michael Covel and Curtis Faith,[16] among others.

The method is based on an *N*-day breakout, with a number of add-on rules that seem to be rooted in trader experience. The following summary should be sufficient to test the concepts, although it is not clear how much discretion was allowing in following the signals. There are a number of subtle rules and variations presented in the original material that are not included here.

During the early 1980s the futures markets were not as diverse as now, and commodities made up a larger part of the portfolio. Only markets that traded on U.S. exchanges were included.

---

[16]Michael W. Covel, *The Complete Turtle Trader* (New York: Collins, 2007); and Curtis Faith, *The Original Turtle Trading Rules* (2003), originalturtles.org.

The Turtle Rules

There are two systems, *S1* and *S2*; however, there is no indication of what proportion of assets are allocated to each. We will assume that the capital is apportioned equally. The two versions have differences in both the positioning of stop-losses and the method of compounding, and those rules are combined here. Both are explained.

System 1 (S1)

1. *Enter a long position* when the intraday high exceeds the daily highs of the previous 20 days. *Exit the long position* when the intraday low falls below the lowest low of the previous 10 days.

2. *Enter a short position* when the intraday low falls below the daily lows of the previous 20 days. *Exit the short position* when the intraday high exceeds the highest high of the previous 10 days.

3. *Filter Rule.* Ignore the *S1* entry if the previous *S1* entry was profitable (whether or not it was taken). However, if the prior *S1* trade was a loss of at least 2*L*, then the trade could be taken. *L* is a 20-day *average-off true range* volatility measure,

$$L_t = \frac{\left(19 \times L_{t-1} + TR_t\right)}{20} \times Conversion \ factor$$

where *TR* is the daily true range, *t* is today, and the *conversion factor* is the amount that transforms the point value of a market into its dollar value. We will call this the *big point value* (*BPV*).

System 2 (S2)

1. *Enter a long position* when the intraday high exceeds the daily highs of the previous 55 days. *Exit the long position* when the intraday low falls below the lowest low of the previous 20 days. Traders were allowed discretion in selecting a period slightly shorter than 55 days.

2. *Enter a short position* when the intraday low falls below the daily lows of the previous 55 days. *Exit the short position* when the intraday high exceeds the highest high of the previous 20 days.

3. There was no filter for method *S2*.

Risk Control

1. *A stop-loss* is placed a distance of 2*L* from the initial entry.

2. Trades are exited on the first occurrence of
   a. The stop-loss
   b. An *S1* or *S2* reversal
   c. A loss of 2% relative to the portfolio (where 2*L* is equal to 2% of the portfolio)

3. *The position size* is determined by equalizing $L \times BPV$ across all markets; then *1 unit size* = (1% *investment*)/($L \times BPV$).

4. *Position limits* restrict the maximum *unit size* that could be traded in correlated markets:

   a. Single market, $4 \times$ unit size

   b. Closely correlated markets (e.g., energy complex, precious metals, currencies, short-term rates), $6 \times$ unit size

   c. Less correlated markets (related by inflation or other factors), $10 \times$ unit size

   d. Any net position, long or short, $12 \times$ unit size

5. *Compounding positions.*

   a. Add another unit (or ½ unit) for every profitable price move equal to L, measured from the actual entry price.

   b. A maximum of 5 units is allowed (this is different from 4a).

   c. A stop-loss is set at ½L on day 1 and $2L$ afterwards (see *alternate stops* below). Once a second unit is bought, all stops are increased to $2L$. For each subsequent unit, stops were brought to $2L$ measured from the most recent entry, so the total trade risk was $2L \times total\ contracts$ at all time during the trade.

   d. *Alternate stops*. Stops could be placed at ½L instead of $2L$. If stopped out, that position could be reentered at the original entry price.

6. *Portfolio risk management.* For every 10% drawdown in the portfolio, measured from the peak equity, the position size was cut by 20%. (At the time the *de facto* loss limit was 50% for the industry.) For every $6\frac{2}{3}$% recovery, 10% was added back. Costs were included when calculating these values.

Comment

Richard Dennis was reported to have been one of the most successful floor traders during the late 1970s and early 1980s. The industry was surprised when he formed the Turtles but expected great success. It is not known if his previous success was based on using this method or one similar to it. Floor traders usually profit from the bid-asked spread during active trading periods, although Dennis was also said to hold large positions overnight.

It often happens that a good period in the market will encourage you to expand your operation. As Fate would have it, that good period is often followed by a bad one. As Will Rogers said, "Good judgment comes from experience, and a lot of that comes from bad judgment."

The basic breakout concept, with the initial position size and stop-loss, was programmed as *TSM Turtle*, available on the Companion Website. While it does not attempt to duplicate all the rules, it is good to see the underlying method returns profits, and subtle additions may add to those returns. The period prior to 1984 should be of interest because it was this period in which Dennis made the decision to form the Turtles.

Copper 20 x 10
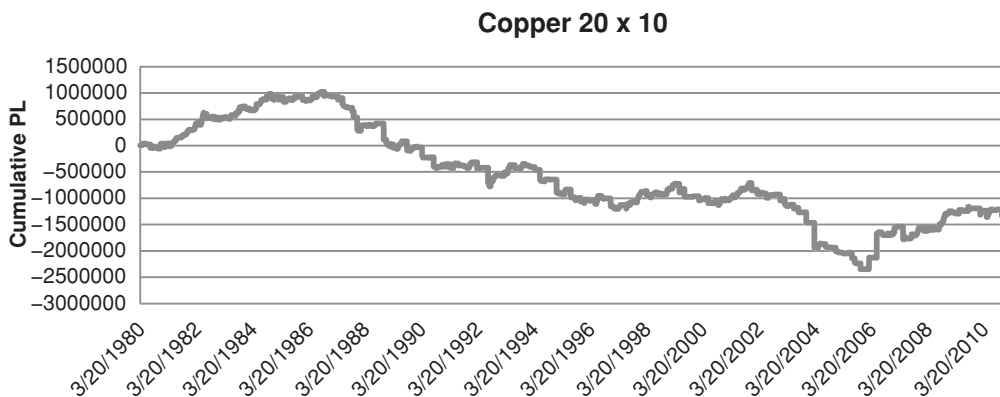
FIGURE 5.29 Cumulative profits for basic Turtle strategy for copper futures using faster parameters.

Copper 55 x 20
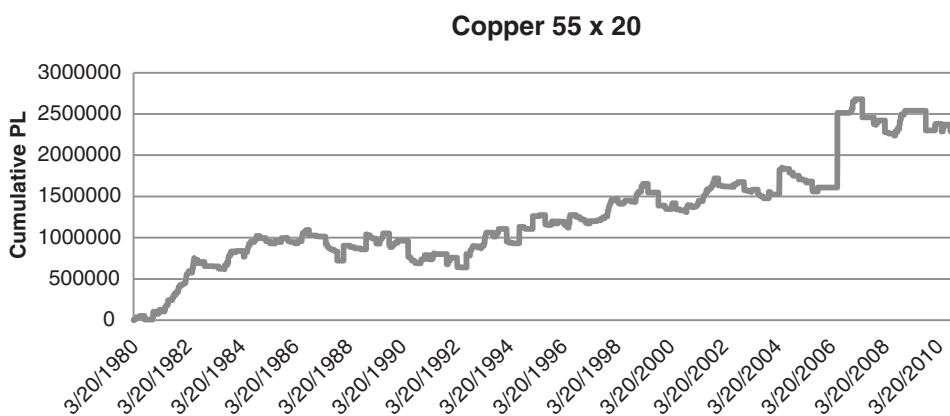
FIGURE 5.30 Cumulative profits for Turtle strategy using slower parameters.

Figures 5.29 and 5.30 show the cumulative profits (closed out trades only) for copper beginning in 1980. The slower system *S2* shows steady profits for the entire 30 years, which is typical of a long-term-trend–following method. The faster system *S1* shows profits only for the first part of the 1980s, the period in which Dennis would have decided to form the Turtles. Remembering that this is only one market and many of the rules are not implemented; the reader would need to look further to decide if this is the general pattern or simply a poor example.
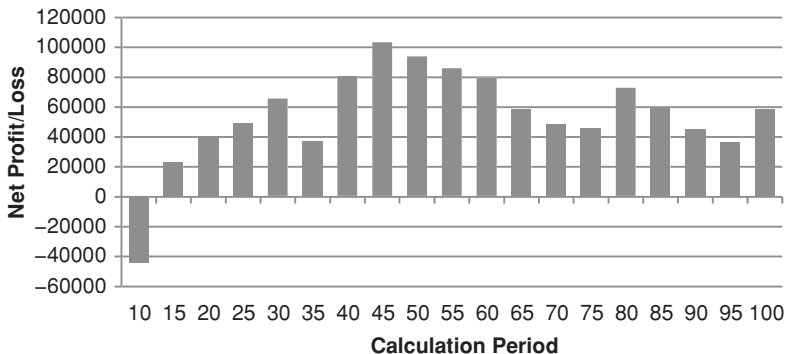
FIGURE 5.31 Relative performance of *N*-day breakout, copper futures, 1980–2010, with calculation periods from 10 to 100 days.

By way of comparison, the same 30-year period was tested using the *N*-Day Breakout and the results of the optimization are shown in Figure 5.31. It shows that copper was a very trending market, and all but one calculation period returned a net profit. In retrospect, the periods of 55 and 20 were not the best choice; however, Dennis must be credited for having the foresight to use breakout systems ahead of the crowd.

Regression Analysis

Regression analysis is a way of measuring the relationship between two or more sets of data. A stock analyst might want to know the relationship between the price of gold per ounce and the share price of Barrick Gold Corporation (ABX). An economist might want to know how the more complex relationship between interest rates, inflation, and the trade balance changes the value of the U.S. dollar. A hedger or arbitrageur could use the relationship to establish the relative fair value of two related products, such as palm oil and soybean oil, in order to select the cheaper product or to profit from price distortions. An investor might simply want to find the strongest stock in the banking sector. Or, you may simply want to find the best straight line fit through a series of prices. Regression analysis involves statistical measurements to determine the type of relationship that exists between the data studied. Many of the concepts are important in technical analysis and should be understood by all technicians, even if they are not used frequently. The techniques may also be used to trade, as is shown at the end of this chapter.

COMPONENTS OF A TIME SERIES

Regression analysis is a tool often applied to identify the main component of a time series, the *trend*. With some variation, it can also be used to isolate the *seasonal* (or secular trend) and *cyclic* components. These three factors are present in all price data. The part of the data that cannot be explained by these three elements is considered *random*, or unaccountable price movement.

Trends are the basis of many trading systems. *Long-term trends* can be related to economic factors, such as changing interest rates, inflation, shifts in the value of the U.S. dollar due to the balance of trade, and even consumer confidence. The reasons for the existence of *short-term trends* are not always clear. Expectations of a merger or government approval of a new drug, a temporary disruption in oil supply, or a dockworkers slowdown could all be catalysts for higher prices. However, trends that survive for only

a few days or weeks cannot be explained by macroeconomic factors, but are usually the result of investor behavior—reaction to the constant flow of minor news and market reports.

Major sustained fluctuations above and below the long-term trend are attributed to *cycles.* Both business and industrial cycles respond slowly to changes in supply and demand. The decision to close a factory or build more container ships or office buildings is not made immediately, nor can the decision be easily changed once it has been made. Stimulating economic growth by lowering interest rates is not a cure that works overnight. Opening a new mine, finding new crude oil deposits, or building an automobile assembly plant in another country makes the response to increased demand slower than the act of cutting back on production. Moreover, once the investment has been made, business is not inclined to stop production, even at marginal returns.

Seasonality, the third component of price movement, is a form of cycle that depends on the calendar year. Most directly affected are crops, which have planting and harvesting cycles attributed directly to Mother Nature. The travel industry is much more active in the summer than winter, and there is a much higher demand for electricity in the summer. Oil refiners shift their refining from heating oil to gasoline in February and March as winter demand comes to an end, then change back again in the late summer. The fashion industry anticipates the changing seasons in order to introduce their new lines of clothing.

The *random* element of price movement is a composite of everything unexplainable. In later sections an Autoregressive Integrated Moving Average (ARIMA), or a Box-Jenkins method, will be used to find shorter trends and cycles that may exist in the leftover data. This chapter concentrates on trend identification, using the methods of regression analysis. Seasonality and cycles are discussed in Chapters 8 and 9. Because the basis of a strong trading strategy is its foundation in real phenomena, traders and serious students of price movement should understand the tools of regression analysis to avoid incorporating erroneous relationships into their strategies. A sound premise is the basis for success.

CHARACTERISTICS OF THE PRICE DATA

A time series is not just a series of numbers, but *ordered pairs* of price and time. There is a special relationship in the way price moves over various time intervals; the way price reacts to periodic reports; and the way prices fluctuate due to the time of year. Most trading strategies use one price per day, usually the closing price, although some methods average the high, low, and closing prices. Economic analysis operates on weekly or monthly average data, but might use a single price (e.g., "week on Friday") for convenience. Two reasons for the infrequent data (also called *low frequency data*) are the availability of most major statistics such as earnings, consumer confidence, or supply and demand, and the intrinsic long-term perspective of the analysis. The use of less frequent data causes a

smoothing effect. The highest and lowest prices will usually not appear on the last day of a month or in average monthly data, and the data will seem less volatile. Even when using daily closing price data, the intraday highs and lows have been eliminated, and the closing prices show less erratic movement.

Selection of the Calculation Period

A regression analysis, which can identify price direction over a specific time period, will not be influenced by cyclic patterns or short-term trends that are the same length as the time interval used in the analysis. For example, if we use 12 months of data, then the average of those 12 months includes an entire season of price changes. If the pattern was a perfect cycle, increasing by 15% during high-demand and dropping by 15% during harvest or high-supply, then the average would neutralize both up and down swings. Therefore, the use of calculation periods equal to multiples of a calendar year is used for *deseasonalizing* the data.

In Figure 6.1, a regression line is calculated using the cash corn prices for the past 30 years. Although it is greatly influenced by the exceptional rally in 2008, the center line increased from $2.36/bu to $2.88 over the 30 years, or 0.73% per year. We can think of that as the inflation rate for corn, which is based on real inflation, the value of the U.S. dollar, improvements in technology that increase the crop size, consumption, and other factors.

A 12-month moving average is also plotted in Figure 6.1 This is also not influenced by the seasonal price changes, but when prices rise over more than one year, an upward trend can be seen. If you do not want to remove the effects of seasonality or longer cycles, the time interval of the regression analysis should be 3 months or less. In this way, a regression analysis, or any trend technique, may be used to identify a seasonal or cyclic element.
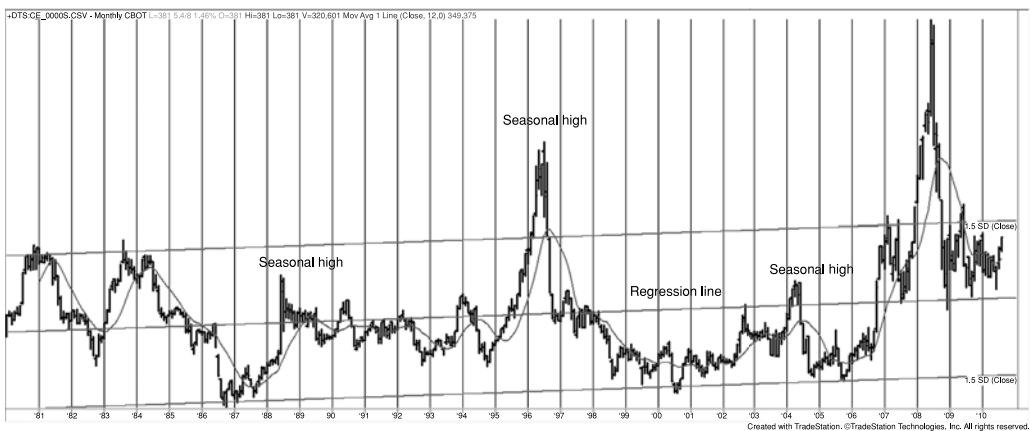
FIGURE 6.1 A basic regression analysis results in a straight line through the center of prices.

In classic time-series decomposition, the trend is removed first, and the detrended series is then used to find the cyclic and seasonal components. This can be done by

1. Subtracting the daily prices from the corresponding daily value of the regression line, or
2. Taking the first differences, $P_t - P_{t-1}$, for the entire data series.

LINEAR REGRESSION

When most people talk about regression, they normally think about drawing a straight line through the center of some period of price movement, such as in Figure 6.1. At the end of this chapter, and in the discussions and comparisons of trend systems, we will use the slope of the linear regression applied to a single price series to create a trend trading method. But regression is a simple and powerful tool for explaining the relationship between two price series. In econometrics, analysts use *multiple regression* to find the relationship between various factors, such as supply, demand, inflation, and price.

In this chapter we look at how regression can be used

- As a trading tool for a single market
- To find the relationships between two markets or two data series
- To rank markets

The inputs to a regression analysis must be two time series of the same frequency, most often daily, weekly, or monthly. For trading we will use daily data, but for an economic relationship monthly is better. We start with the linear relationship between two price series, X and Y. A linear relationship will try to find the value of Y (the *dependent variable*) for each value of X (the *independent variable*) using the formula for a straight line, $Y = a + bX$, where a is the *Y-intercept* of the line, and b is the *slope*, the point where the line crosses the Y axis when the value of X is zero. The linear regression is also called a *straight-line fit*, or simply the *best fit*. It selects the straight line that comes closest to most of the data points. The result tells you that, for example, for every move of $1 in series X, we can expect a move of $1.50 in series Y, or we can expect inflation to increase the price of corn by 0.000029% per day. We will apply a linear regression in two examples, first to explain the price movement of corn based on the price of soybeans, then to explain the price of Barrick Gold Corporation (ABX) based on the price of physical gold.

Explaining, Not Predicting

You may have noticed that we refer to *explaining* the Barrick Gold stock price in terms of the price of gold bullion. We do not say that we can *predict* the price. We are finding the past relationship between two price series. You may decide that this relationship can

be used to trade those two markets whenever prices move too far from one another. The regression analysis may establish what you see as a *fair value* for one market based on the price of the other. In order to forecast a price, you will need to establish that conditions at the date of your forecast are likely to be the same as the period over which the regression was calculated. You must also account for the loss of accuracy, or confidence, as you forecast further into the future.

Calculating the Best Straight-Line Fit

The simplest example of linear regression, called the *method of least squares*, is the one used most often by traders, and the one seen in Figure 6.1, calculating the best straight-line fit through a selected period of price movement. This is done in the same way as finding the relationship between two price series, except that we will substitute the simple sequence 1, 2, 3, 4, . . . for the second series. We do not use the date as the dependent variable because weekends and holidays break the continuity of the series.

We choose 10 days of price movement in Wal-Mart during March 2002. In order to find the best straight-line fit, we begin with the equation for a straight line

$$Y = a + bX$$

In this equation, Y is the dependent variable because it is a function of the value of X, the independent variable. The slope, b, is the relative change in Y for every unit change in X. Therefore, if $b = 0.20$ and X are sequential days, then for every day, the price (Y) gains $0.20. The Y-intercept, a, is an adjustment in the price level to bring X and Y into alignment. It is also the point at which the straight line crosses the Y-axis when $X = 0$.

Method of Least Squares

The *method of least squares* finds the straight line that comes closest to all prices. To do this, calculate the *sum of the squares* of all the differences between the price and the corresponding value of the straight line and choose the line that has the smallest total deviation. The mathematical expression for this is

$$S = \sum_{i=t-n+1}^{t} \left(y_i - \hat{y}_i \right)^2$$

where S = the sum of the squares of the error at each of the 10 Wal-Mart points on the straight line (one point for each price, designated by $i = 1, 2, 3, \ldots$)

                y_i = the price of Wal-Mart on day i

                $\hat{y}_i$ = the estimated value of this price on the straight line

      $y_i - \hat{y}_i$ = the difference between the actual price y_i at i and the predicted price $\hat{y}_i$

Graphically, the individual deviations, or errors, for the first four points may look like those in Figure 6.2. Each actual data point is $(1, y_1)$, $(2, y_2)$, $(3, y_3)$, . . . and the
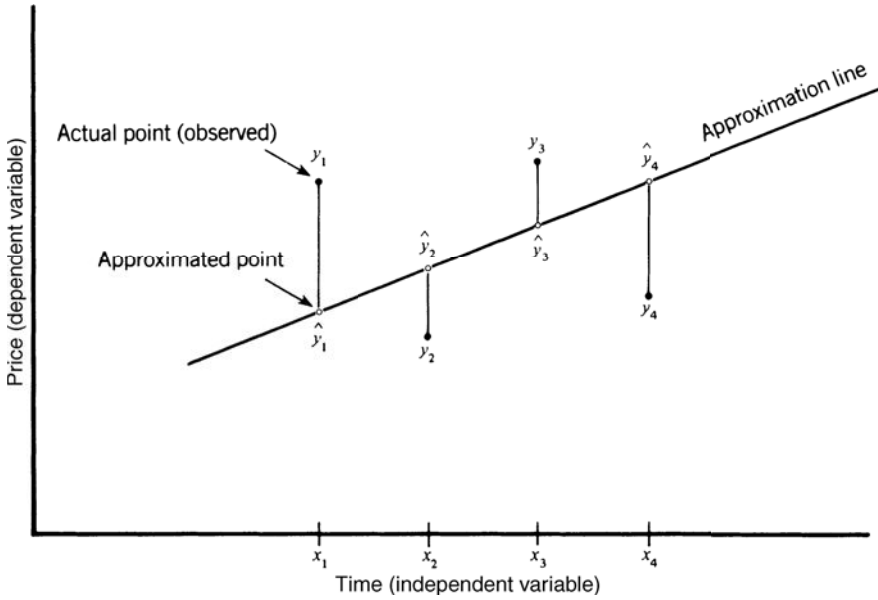
FIGURE 6.2   Error deviation for method of least squares.

corresponding position on the regression line is $(1, \hat{y}_1)$, $(1, \hat{y}_2)$, $(1, \hat{y}_3)$, ... The sum of the squares of the errors is

$$S = \sum_{i=t-n+1}^{t} \left(y_i - \hat{y}_i\right)^2$$

$$= \left(y_1 - \hat{y}_1\right)^2 + \left(y_2 - \hat{y}_2\right)^2 + \left(y_3 - \hat{y}_3\right)^2 + \left(y_4 - \hat{y}_4\right)^2$$

The straight line that causes S to be the smallest possible value will be the best choice for these data points. The square of $y_i - \hat{y}_i$ is always positive, therefore, prices above and below the regression line are treated equally.

To use the least-squares method for solving the Wal-Mart/time-price relationship, look for the solution to the straight line, $Y = a + bX$, expressed as

$$b = \frac{N\sum xy - \sum x \sum y}{N\sum x^2 - \left(\sum x\right)^2}$$

$$a = \frac{1}{N}\left(\sum y - b\sum x\right)$$

where $N =$ the number of data points (10 in the Wal-Mart example)
　　　　$\Sigma =$ the sum over N points

TABLE 6.1 Calculations for a Wal-Mart Least-Squares Fit

Date	Wal-Mart Price (Y)	Sequence Number (X)	X²	XY	Y²	Regression Value
9/21/2001	44.66	1	1	44.66	1994.52	46.25
9/24/2001	47.28	2	4	94.56	2235.40	46.96
9/25/2001	48.40	3	9	145.20	2342.56	47.66
9/26/2001	49.48	4	16	197.92	2448.27	48.37
9/27/2001	49.39	5	25	246.95	2439.37	49.08
9/28/2001	49.50	6	36	297.00	2450.25	49.78
10/1/2001	49.76	7	49	348.32	2476.06	50.49
10/2/2001	51.60	8	64	412.80	2662.56	51.20
10/3/2001	52.73	9	81	474.57	2780.45	51.90
10/4/2001	51.50	10	100	515.00	2652.25	52.61
Sums	494.30	55	385	2776.98	24481.69	

In order to solve these equations, use a spreadsheet and enter the Wal-Mart prices in column B (see Table 6.1). Construct a table where all of the individual expressions in the two formulas can be calculated (see Table 6.1).[1] We can now substitute the values on the *Sums* line into the two equations:

$$b = \frac{10(2776.98) - (55)(494.30)}{10(385) - (55)(55)}$$
$$= \frac{27769.8 - 27186.5}{3850 - 3025}$$
$$= 0.70703$$
$$a = \frac{1}{10}(494.30 - 70703 \times 55)$$
$$= \frac{455.413}{10}$$
$$= 45.5413$$

The equation for the least-squares approximation is

$$Y = 45.5413 + 0.70703X$$

Selecting values for X and solving for Y gives the results shown in the right column of Table 6.1, representing the straight line fit. This is shown along with the original prices for Wal-Mart in Figure 6.3. The straight line approximation increases by 0.707¢ per day and the approximation line starts at 45.54, where $X = 0$.

---

[1] The Companion Website (www.wiley.com/go/tradingsystems5e) has this spreadsheet in a form that can be used for other calculations.
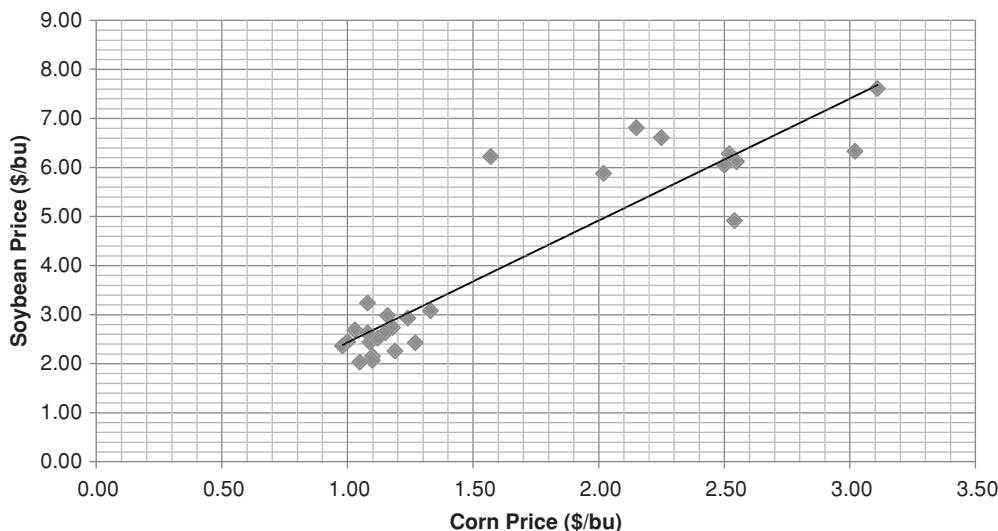
FIGURE 6.3   Scatter diagram of corn-soybean pairs with linear regression solution.

Excel Regression

Excel makes this much easier, although the formulas given above are easily programmed directly if you are not using Excel. Excel's regression function can be found in the drop-down menu Data/Data Analysis (this is available as an "add-in" if not already loaded), then "Regression." Select the sequential values as *X*, the price as *Y*, and the output any cell to the right of the data. Excel will produce the analysis shown in Table 6.2. The *Y*-intercept and slope can be found in the bottom left of the table. Also of interest is the correlation, $R = 0.92$, at the top, which shows the quality of the fit (which is very good).

Corn Explained by Soybeans

Finding the relationship between corn and soybean prices will tell the farmer whether planting one or the other is a better business decision. We use the same technique, but this time we use the corn and soybean price series instead of letting one be the simple integer sequence. Table 6.3 gives the cash prices and the intermediate calculations. These result in the slope $= 0.282$ and the *y*-intercept $= 0.336$, giving the equation for the regression as $y = 0.282 + 0.336x$.

The Excel solution is shown in Table 6.4 with the correlation, $R = 0.91$. Looking at the scatter diagram of the corn-soybean relationship, Figure 6.3, we see that if corn is selling at $2.00/bu we can expect soybeans to sell at $5.00/bu.

This is just what would be expected for farm income. Because the corn yield per acre is 2.5 times greater than the soybean yield in most parts of the United States, the ratio 1:2.5 will yield $5 for soybeans when corn is $2.

TABLE 6.2 Output from Excel's Regression Function

Summary Output

Regression Statistics

Multiple R	0.923
R Square	0.851
Adjusted R Square	0.833
Standard Error	0.948
Observations	10.000

ANOVA

	df	SS	MS	F	Significance F
Regression	1.000	41.241	41.241	45.843	0.000
Residual	8.000	7.197	0.900		
Total	9.000	48.438			

	Coefficients	Standard Error	t Stat	P-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	45.541	0.648	70.287	0.000	44.047	47.035	44.047	47.035
X Variable 1	0.707	0.104	6.771	0.000	0.466	0.948	0.466	0.948

TABLE 6.3 Intermediate Calculations for the Corn-Soybean Regression

Year	Sequence	Corn y_i	Soybeans x_i	x_i^2	x_iy_i	y_i^2
1956	1	1.27	2.43	5.90	3.09	1.61
1957	2	1.19	2.26	5.11	2.69	1.42
1958	3	1.10	2.15	4.62	2.36	1.21
1959	4	1.10	2.07	4.28	2.28	1.21
1960	5	1.05	2.03	4.12	2.13	1.10
1961	6	1.00	2.45	6.00	2.45	1.00
1962	7	0.98	2.36	5.57	2.31	0.96
1963	8	1.09	2.44	5.95	2.66	1.19
1964	9	1.12	2.52	6.35	2.82	1.25
1965	10	1.18	2.74	7.51	3.23	1.39
1966	11	1.16	2.98	8.88	3.46	1.34
1967	12	1.24	2.93	8.58	3.63	1.54
1968	13	1.03	2.69	7.24	2.77	1.06
1969	14	1.08	2.63	6.92	2.84	1.17
1970	15	1.15	2.63	6.92	3.02	1.32
1971	16	1.33	3.08	9.49	4.10	1.77
1972	17	1.08	3.24	10.50	3.50	1.17
1973	18	1.57	6.22	38.69	9.76	2.46
1974	19	2.55	6.12	37.45	15.61	6.50
1975	20	3.02	6.33	40.07	19.12	9.12
1976	21	2.54	4.92	24.21	12.50	6.45
1977	22	2.15	6.81	46.38	14.64	4.62
1978	23	2.02	5.88	34.57	11.88	4.08
1979	24	2.25	6.61	43.69	14.87	5.06
1980	25	2.52	6.28	39.44	15.83	6.35
1981	26	3.11	7.61	57.91	23.67	9.67
1982	27	2.50	6.05	36.60	15.13	6.25
Sums		Σ 43.38	106.46	512.95	202.35	82.27

Data source: Illinois Statistical Service; 1966–1982, Commodity Research Bureau *Commodity Yearbook*.

Gold Bullion and Barrick Gold Corporation

A question often asked is "If gold goes up by $1, how much will a gold mining share go up?" For this to make sense, the mining company should not be diversified but concentrated on gold production. To find how the stock price of Barrick Gold Corporation can be explained by the price of physical gold, we will look only at the six months following the tragedy of September 11, 2001. If you have reviewed the method of solving the corn-soybean relationship, you will understand how to interpret the results of the spreadsheet solution.

First, prepare the data for the Excel solution. We import the prices of ABX and gold into a worksheet and align the dates. Fill data forward for any missing dates. Next, index each of the price series, beginning with 100. This step is not essential for this solution, but

TABLE 6.4 Excel Solution for the Corn-Soybean Regression

Summary Output

Regression Statistics

Multiple R	0.9135
R Square	0.8345
Adjusted R Square	0.8279
Standard Error	0.2888
Observations	27.0000

ANOVA

	df	SS	MS	F	Significance F
Regression	1.0000	10.5106	10.5106	126.0481	0.0000
Residual	25.0000	2.0846	0.0834		
Total	26.0000	12.5952			

	Coefficients	Standard Error	t Stat	P-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.2825	0.1304	2.1665	0.0400	0.0139	0.5510	0.0139	0.5510
X Variable 1	0.3358	0.0299	11.2271	0.0000	0.2742	0.3974	0.2742	0.3974

TABLE 6.5   Setting Up a Spreadsheet for Calculating a Regression Solution

| Date | Actual Prices | | Indexed Prices | | Estimated | |
	ABX	Gold	ABX	Gold	From Actual ABX	From %ABX
9/10/2001	16.11	271.50	100.00	100.00	15.90	98.68
9/17/2001	16.65	293.20	103.35	107.99	17.20	106.77
9/18/2001	16.70	289.40	103.66	106.59	16.97	105.36
9/19/2001	16.80	288.10	104.28	106.11	16.90	104.87
9/20/2001	17.35	288.50	107.70	106.26	16.92	105.02
9/21/2001	17.17	292.50	106.58	107.73	17.16	106.51
9/24/2001	16.67	288.20	103.48	106.15	16.90	104.91
9/25/2001	16.32	288.00	101.30	106.08	16.89	104.84
9/26/2001	17.34	291.40	107.64	107.33	17.09	106.11
9/27/2001	17.62	289.30	109.37	106.56	16.97	105.33
9/28/2001	17.35	293.10	107.70	107.96	17.20	106.74
10/1/2001	17.70	290.60	109.87	107.03	17.05	105.80
10/2/2001	17.95	291.60	111.42	107.40	17.11	106.18
10/3/2001	17.10	290.10	106.15	106.85	17.02	105.62
10/4/2001	17.00	290.30	105.52	106.92	17.03	105.69
10/5/2001	17.25	291.20	107.08	107.26	17.08	106.04
10/8/2001	16.86	291.80	104.66	107.48	17.12	106.26
10/9/2001	16.44	290.00	102.05	106.81	17.01	105.58
10/10/2001	15.95	287.10	99.01	105.75	16.84	104.51

it will be necessary when comparing markets for ranking (a later section in this chapter) and to get the results in percentages. In this section, the relationship will be solved both with the original data and with the indexed data. Each index value is $I_t = I_{t-1} \times p_t / p_{t-1}$, where p_t is today's price (see Chapter 2). The two index series now are in the same percentage notation, as seen in Table 6.5.

Using the Excel regression function applied to the raw prices, the slope is 0.060 and the y-intercept is −0.418. This means that, for every \$1 increase in gold, ABX will increase by 6¢. Using the indexed values, the slope is 1.012 and the y-intercept is −2.586. For every 1% increase in gold, ABX increases by 1.012%, showing that there is essentially a 1:1 relationship between gold and ABX (in this very short example). To calculate the estimated value from the indexed results, the formula is $ABX_t = 1.012 \times gold_t − 2.586$. The estimated value of ABX is shown for both the actual prices and the indexed prices in Figures 6.4a and b. In both cases, the estimated price of ABX tracks gold very closely, yet offering an interesting trading opportunity for arbitraging gold and ABX when they move apart. This will be discussed in Chapter 13.

Programming and Spreadsheet Tools

In addition to the regression function in Excel data analysis, there are simple functions that can produce the same results. *Slope(y, x)* returns the slope value, b, when you enter
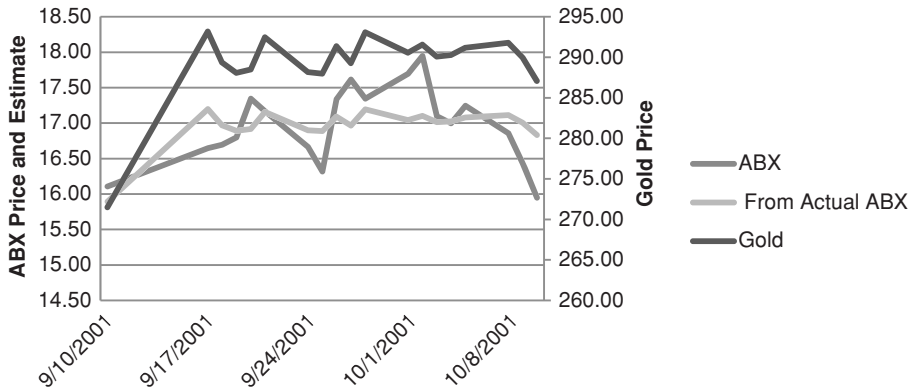
FIGURE 6.4a   Estimated ABX price from actual ABX and gold prices.
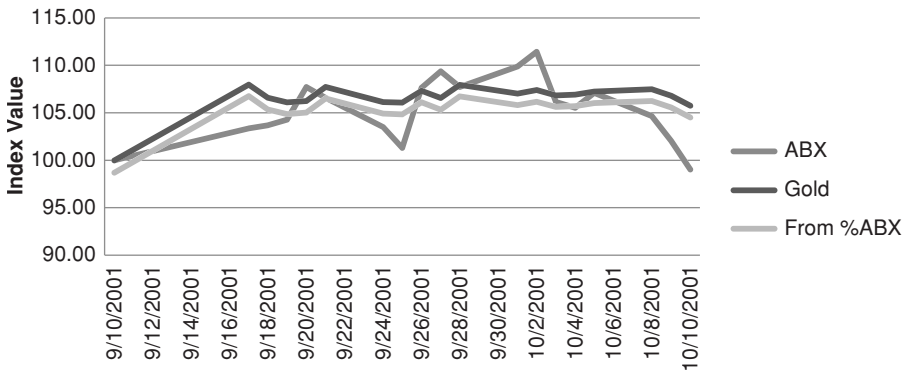
FIGURE 6.4b   Estimated ABX price from indexed ABX and gold prices.

the column lists for *Y* and *X*. *Intercept(y, x)* returns the *y*-intercept, *a*, when you enter the same values for *Y* and *X*. Once you know what to look for, it becomes easy.

In strategy development software, such as TradeStation, the functions that calculate linear regression are nearly the same. *LinearRegSlope* or *LinearRegAngle*, *LinearRegValue, LinRegIntercept*, all produce the estimated values using the dependent variable. Most often, traders are only interested in the slope as a way to identify the direction of prices. Using daily data, they add the slope value to the most recent price to project tomorrow's price,

$$\textit{Projected price (1-day ahead)} = \textit{Today's price} + \textit{Slope}$$

If they want to project *n* days ahead, they multiply the slope by *n*,

$$\textit{Projected price (n-days ahead)} = \textit{Today's price} + n \times \textit{Slope}$$

As prices are projected further ahead, there is a much greater chance of error. This is discussed in a later section, "Confidence Bands."

LINEAR CORRELATION

Solving the least-squares equation for the best fit does not mean that the answer is useful. In the previous sections we used price series that had a clear relationship; therefore, the results appeared valid. The least-squares method will always give an answer, even when there is nothing that makes one prices series dependent upon the other. You may think that two data items affect one another, such as the amount of disposable income and the purchase of television sets, but that might not be the case. Instead, the purchase of television sets may peak just before the Super Bowl.

The *linear correlation*, which produces a value called the *coefficient of determination*, r^2, or the *correlation coefficient*, expresses the strength of the relationship between the data on a scale from +1 (perfect correlation) to 0 (no relationship). It may be even better to look at r, which varies from +1 to −1, rather than r^2, because you then know whether the correlation is positive or negative. A negative correlation between two prices series means that when one goes up, the other goes down. When $r = +1$ there is a perfect positive correlation, when $r = 0$ there is no correlation, and when $r = -1$ there is a perfect negative correlation, as shown in Figure 6.5. This is the most practical way to
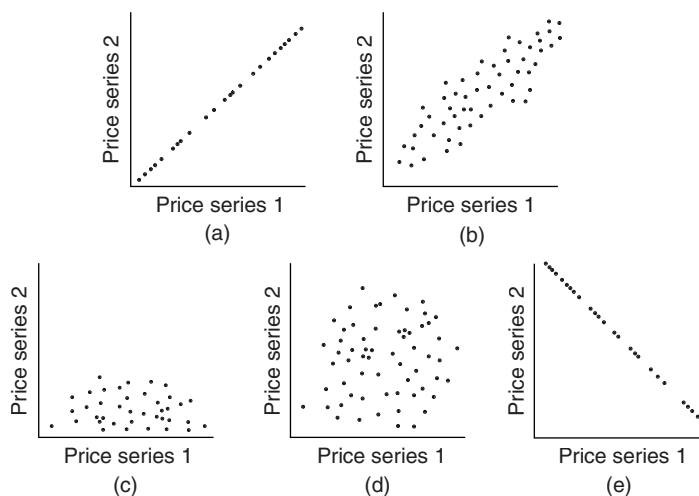
FIGURE 6.5   Degrees of correlation. (a) Perfect positive linear correlation ($r = 1$). (b) Somewhat positive linear correlation ($r = 0.5$). (c) and (d) No correlation ($r = 0$). (e) Perfect negative linear correlation ($r = -1$)

Source: William Mendenhall and James Reinmuth, *Statistics for Management and Economics*, 2nd ed. (Belmont, CA: Wadsworth, 1974). Reprinted by permission.

find out whether two price series are moving in a similar manner. If r^2 is less than about 0.20, then the linear regression has no practical value. In the examples that looked at the corn-soybean relationship, the Excel results showed an $r^2 = 0.83$, very high, while the ABX-gold correlation, whether using prices or indexing, was only 0.26. For arbitrage purposes, however, it is more important to look at *cointegration* rather than correlation, and this is covered in Chapter 13.

To calculate the correlation coefficient, r, for the most recent n days, where t is today, we use *Pearson's Product-Momentum Correlation*

$$r = \frac{1}{N} \sum_{i=t-n+1}^{t} \left[\left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right) \right]$$

In the preceding calculation, s_x and s_y are the standard deviations of x and y, and $\bar{x}$ and $\bar{y}$ are the averages. The result r is interpreted as follows:

$r = +1$	A perfect positive linear correlation. The data points are along a straight line going upward to the right (Figure 6.5a). For every upwards move in x there is a corresponding upwards move in y.
$+1 > r > 0$	The scattered points become more uniformly distributed about a positive approximation line as the value of r becomes closer to +1 (Figure 6.5b).
$r = 0$	No linear correlation exists (Figures 6.5c, 6.5d).
$-1 < r < 0$	The scattered points become more uniformly distributed about a negative approximation line as the value of r becomes closer to –1.
$r = -1$	A perfect negative linear correlation, the line going downward to the right (Figure 6.5e). For every upwards move in x there is a corresponding downwards move in y.

In Excel, the correlation function is *correl(x,y)*, which returns r^2. The formula appears to be different but it is only a regrouping of the factors. If you need to know if the correlation is positive or negative, it will be necessary to do the calculation yourself. The Companion Website spreadsheet, *TSM ABX-Gold regression comparison of power fits*, compares the formula given above with the Excel method.

Correlation Adjustments When Using a Time Series

Because most price analyses involve the use of two time series, precautions should be taken to avoid a trend that can distort the results. A long-term upward or downward trend will overshadow the smaller movements around the trend and exaggerate the correlation. Inputs to correlations are typically price changes, not the actual prices. It is best if these changes are in percent, but for back-adjusted futures data, the price differences should be used. Most software, such as Excel, will calculate the correlation on whatever series you input, so it is your responsibility to detrend it first.

Forecasting Using Regression

A distinct advantage of regression analysis is that it allows the analyst to forecast price movement. In the case of the linear regression, the forecast is simply an extension of the straight line. If you consider that the regression values give the mean over the calculation period, then the forecast is an extension of the mean, which, if you have no other compelling information, is always the safest forecast. Later in this chapter, there will be other *nonlinear* solutions that are used to predict more complex patterns.

The regression forecast is the basis for the *probabilistic model*. Instead of the corn-soybean relationship, find the slope and *Y*-intercept of a straight line fit using the price of soybeans and sequential time as we did in the first Wal-Mart example, Table 6.2. The result is

$$Y = 0.987 + 0.221X$$

where   *Y* = the price of soybeans for the year *X*

Because the solution substituted the value of 1 for 1956 and 27 for 1982, the average farm income per bushel of soybeans is forecast as 7.32 in 1985 ($X = 30$) and 8.38 in 1990 ($X = 35$), increasing 22¢ per year. That would put the price of soybeans at \$13 in 2011, where it's now trading at slightly over \$14.

Confidence Bands

Regression analysis includes its own measure of accuracy called *confidence level.* It is based on a probability distribution of the errors in the fitted equation and the size of the data sample. Looking at Figure 6.3, the straight line cannot touch all the points, but its "goodness of fit" may be measured by using the *standard deviation of the errors* to determine the variance over the total number of data points *N*. The *errors*, *e*, are the difference between the actual data points and the corresponding point on the least-squares line. If the actual data points are y_i and their corresponding value on the fitted line $\hat{y}_i$, then *s* is the standard deviation of the errors,

$$s = \sqrt{\frac{\sum e^2}{N}}$$

where each $e_i = y_i - \hat{y}_i$

Referring to the table of normal distribution (Appendix 1), the 95% level is equivalent to 1.96 standard deviations. Then, a *confidence band* of 95%, placed around the forecast line, is written

$$95\% \text{ upper band} = y_i + 1.96\sigma$$

$$95\% \text{ lower band} = y_i - 1.96\sigma$$
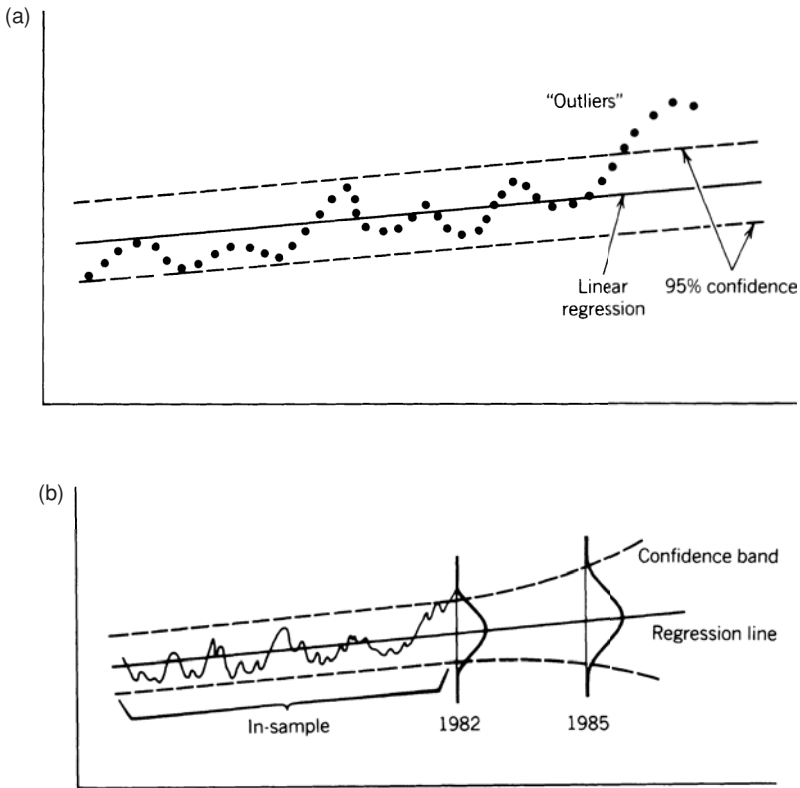
FIGURE 6.6 Confidence bands. (a) Soybeans with 95% confidence band. (b) Out-of-sample forecasts lose confidence.

Figure 6.6a shows the soybean forecast with a 95% confidence band. The points that are outside the band are of particular interest and can be interpreted in one of two ways.

1. They are not representative of normal price behavior and are expected to correct the levels within the bands.
2. The model was not performed on representative or adequate data and should be reestimated.

Figure 6.6b also indicates that the forecast loses accuracy as it is further projected; the forecast is based on the size of the sample used to find the regression coefficients. The more data included in the original solution, the longer the forecast will maintain its accuracy.

In the previous example of the corn-soybean relationship, the Excel analysis output (Table 6.4) showed the values corresponding to the 95% confidence level for both the slope and *Y*-intercept. Generally, the slope is the more interesting value. The solution for the slope was 0.3358 with the *upper 95%* at 0.3974 and the *lower 95%* at 0.2742, 18%

higher and 22% lower respectively. These values are large because only a few data points were used.

NONLINEAR APPROXIMATIONS FOR TWO VARIABLES

The linear regression, $a + bx$, also called a *first-order polynomial* equation, is the simplest way of finding the relationship between two price series. It uses one multiplier (the slope) of power 1 (hence *first* order) and one constant (the *Y*-intercept) to shift the starting point. By adding a third term, cx^2, the approximation can be made much more accurate when a factor, such as the decline in the supply of gold, would cause the stock price to rise at a faster rate. The third term, x^2, introduces a *parabolic* curve—a single, smooth change of direction. The fourth term, dx^3, would add *inflection*, the ability for the curve to turn up and down at different points. Each time we add one more term, the ability to fit one price series based on another becomes better. Because we will want to create a polynomial with more than two or three terms, it is more convenient to change the general notation to

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

The first two terms on the right side of the equal sign form the equation for a straight line (where a_0 was called a and a_1 was called b). For most price forecasting, the *second-order* equation, also called *curvilinear*, which uses three terms, is sufficient. Figure 6.7 shows the general form of a second-order equation, curving up and to the right at an increasing rate.
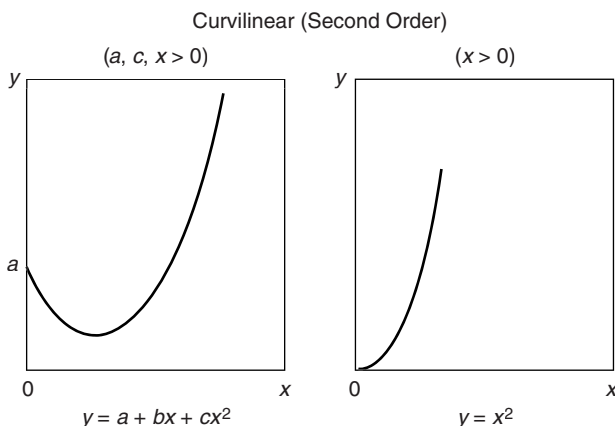
FIGURE 6.7 Curvilinear (parabolas).

Source: Copyright ©1976 From Standard Mathematical Tables, 24th Edition by William H. Beyer. Reproduced by permission of Taylor and Francis Group, LLC, a division of Informa plc.

If we want to know if the price of ABX will go up faster than the linear approximation when gold prices move much higher, then we will want to try a second- and third-order approximation. Instead of a period of only a few days, this solution will use 10 years from 1998. A spreadsheet with this data and the results of the approximations, *TSM ABX-Gold Regression*, is available on the Companion Website.

The second-order (curvilinear) form

$$y = a + bx + cx^2$$

must be solved for the coefficients *a*, *b*, and *c* using the simultaneous equations

$$Na_0 + a_1 \sum x + a_2 \sum x^2 = \sum y$$

$$a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 = \sum xy$$

$$a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 = \sum x^2 y$$

Second-Order Least-Squares Solution

The least-squares solution can be extended to the curvilinear (second-order) equation by minimizing the sum of the errors[2]

$$S = \sum_{i=1}^{N} \left(y_i - a_0 - a_1 x_i - a_2 x_i^2 \right)^2$$

First, it is necessary to separate the various intermediate sums before expressing the solution for a_0, a_1, and a_2.

$$S_{xx} = \frac{1}{N} \sum_{i=1}^{N} \left(x_i - \overline{x} \right)^2$$

$$S_{xy} = \frac{1}{N} \sum_{i=1}^{N} \left(x_i - \overline{x} \right) \left(y_i - \overline{y} \right)$$

$$S_{yy} = \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \overline{y} \right)^2$$

---

[2] F. R. Ruckdeschel, *BASIC Scientific Subroutines*, vol. 1 (Peterborough, NH: Byte/McGraw-Hill, 1981).

$$S_{xx^2} = \frac{1}{N} \sum_{i=1}^{N} \left(x_i - \bar{x} \right) \left(x_i^2 - \bar{x}^2 \right)$$

$$S_{x^2x^2} = \frac{1}{N} \sum_{i=1}^{N} \left(x_i^2 - \bar{x}^2 \right)^2$$

$$S_{yx^2} = \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \bar{y} \right) \left(x_i^2 - \bar{x}^2 \right)$$

The constant values can then be found by substitution into the following equations

$$a_1 = \frac{S_{xy} S_{x^2x^2} - S_{yx^2} S_{xx^2}}{S_{xx} S_{x^2x^2} - \left(S_{xx^2} \right)^2}$$

$$a_2 = \frac{S_{xx} S_{yx^2} - S_{xx^2} S_{xy}}{S_{xx} S_{x^2x^2} - \left(S_{xx^2} \right)^2}$$

$$a_0 = \bar{y} - a_1 \bar{x} - a_2 \bar{x}^2$$

The procedure is identical to the linear least-squares solution. Fortunately, there are simple computer programs that have already been written to solve these problems, such as Polysoftware's *Pro-Stat* or Matlab. Alternatively, Excel can be used if the inputs are transformed.

ABX-Gold Results

The calculations from Table 6.3, including the additional sums for x^3, x^4, and x^2y, can be substituted into the preceding equations. For the general solution, the system of simultaneous linear equations can be solved by the process of matrix elimination, a technique that should not be performed without the help of a computer. However, for those foolish enough to try, an example can be found in the Appendix.[3] For this solution we'll use the Pro-Stat software, which imports the data, asks for the independent and dependent variables, then the order (2, 3, 4, …), then gives a solution which is summarized in Table 6.6. For a fully programmable, integrated platform, Matlab is the preferred choice.

Figures 6.8 and 6.9 give a visual interpretation of the ABX-gold relationship. Figure 6.8 overlays the gold and ABX prices on the same chart with the ABX prices scale on the right and cash gold on the left. For the first eight years, from 1998 through 2006, prices tracked each other very closely. It was only as volatility became extreme during the most recent years that the pattern varied, mostly in different volatility than in actual direction.

The results of the linear, second-order, and third-order solutions are shown in Figure 6.9. The linear and third-order solutions track the price of ABX very closely up

---

[3] See Appendix 2 for examples of matrix solutions.

TABLE 6.6	ABX = f(gold) Solution for 1st-, 2nd-, and 3rd-Order Polynomials			
	a_0	a_1	a_2	a_3
Linear	5.0038	0.036955		
2nd Order	2.6391	0.046978	8.74E-06	
3rd Order	−0.2895	0.065192	−4.27E-05	8.74E-06
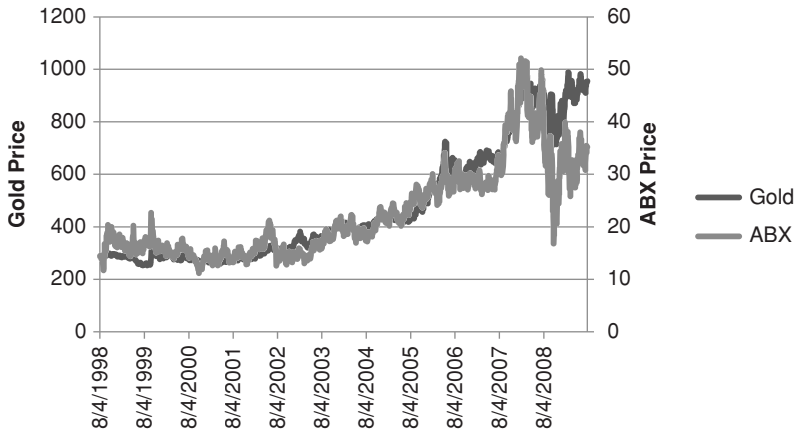
FIGURE 6.8 ABX and cash gold prices, showing different levels of volatility in more recent years.
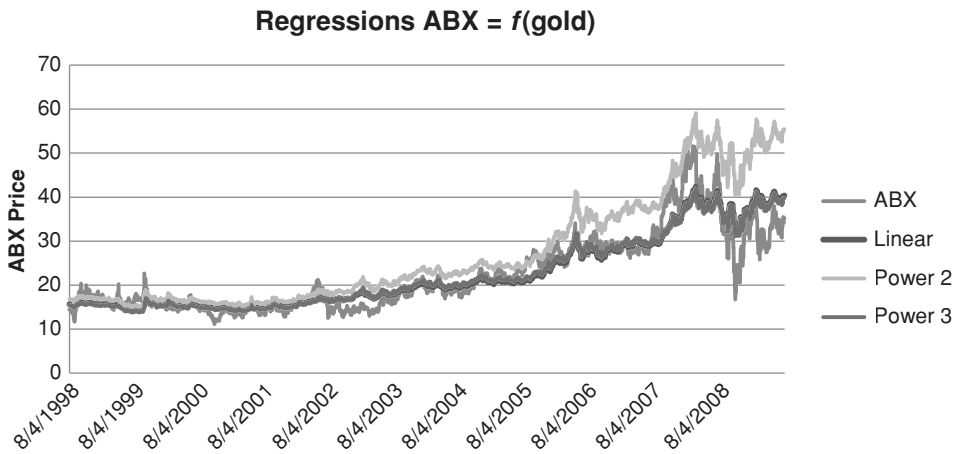
FIGURE 6.9 Comparison of 2nd-order and 3rd-order fits for ABX-gold.

until volatility increases in late 2006. The pattern is right, but the forecasted ABX prices are much less volatile, which is understandable given that more data exists with low volatility. If prices return to pre-2006 volatility levels, the fit should again be good. It is interesting that the second-order solution had a better volatility fit but forecasted prices to be much higher given higher volatility. A simple calculation of the variance of the forecast from the actual price would show which solution is best.

TRANSFORMING NONLINEAR TO LINEAR

Two curves that are often used to forecast prices are *logarithmic* (*power*) and *exponential* (see Figure 6.10). The exponential, curving up, is used to scale price data that become more
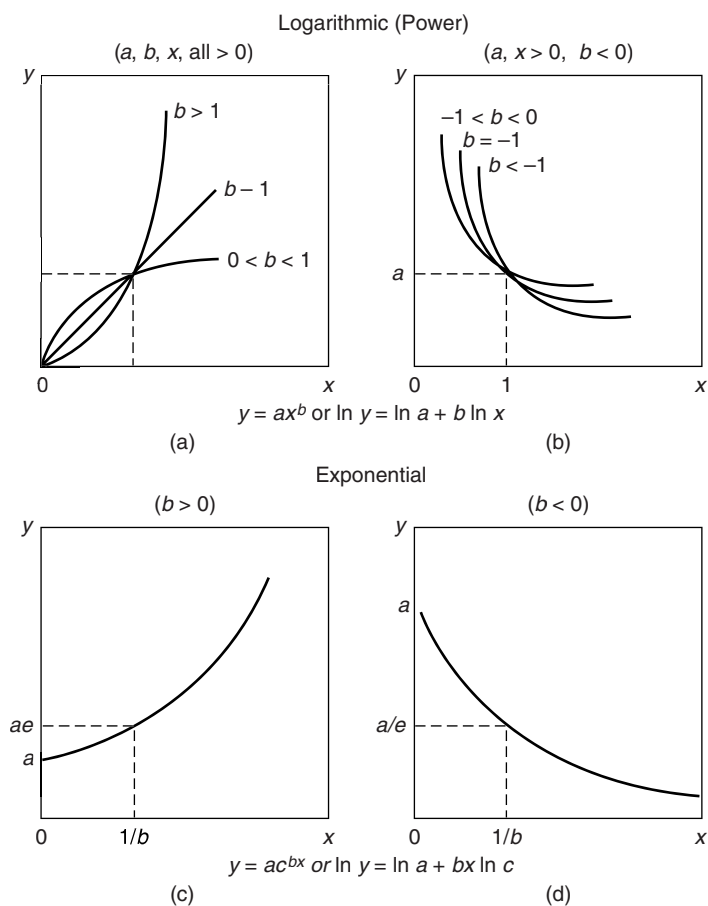
FIGURE 6.10   Logarithmic and exponential curves.

Source: Cuthbert Daniel and Fred S. Wood, *Fitting Equations to Data: Computer Analysis of Multifactor Data,* 2nd ed. (New York: John Wiley & Sons, 1980), 20, 21. Reprinted with permission of John Wiley & Sons, Inc.

TABLE 6.7	Log and Exponential Values for Corn and Soybeans							
Soybeans	x	1.00	2.00	3.00	4.00	5.00	6.00	7.00
Corn (log)	y	0.54	0.94	1.30	1.64	1.96	2.26	2.56
(exp)	y	0.84	1.02	1.24	1.50	1.83	2.22	2.70

volatile at higher levels. The curve that bends down ($0 < b < 1$ in Figure 6.10a) is typical of volatility measured over increasing time intervals. Each of these forms can be easily transformed into linear relationships and solved using the method of least squares. This will allow you to fool the computer into solving a nonlinear problem using a linear regression tool.

To solve the logarithmic relationship, substitute ln x for x, ln a for a, and ln y for y. Instead of using a column of x-values in the spreadsheet, create a new column that has the values ln x. Create another column with ln y. Now solve the regression using ln x and ln y instead of x and y. When completed, you will have the value of b, but you will also have ln a instead of a. The value of ln a must be raised to the power of e to restore the original value of a, $a = (\ln a)^e$. The Excel function for the power of e is *exp*. Once you have the values a and b, the logarithmic approximation can be calculated and plotted.

The significant difference in the log and exponential transformations is that the value of x is not scaled for the exponential. Taking the original corn and soybean data and performing the natural log functions *ln* for both log and exponential approximations results in the linear form that can be solved using least squares. Returning to the simpler corn-soybean example, selected results are shown in Table 6.7.

EVALUATION OF TWO-VARIABLE TECHNIQUES

Of the three curve-fitting techniques, the curvilinear and exponential results are very similar, both curving upward and passing through the main cluster of data points at about the same incline. The log approximation curves downward after passing through the main group of data points at about the same place as the other approximations. To evaluate objectively whether any of the nonlinear methods are a better fit than the linear approximation, find the standard deviation of the errors, which gives a statistical measurement of how close the fitted line comes to the original data points. The results show that the curvilinear is best; the logarithmic, which curves downward, is noticeably the worst (see Figure 6.11). When we display system performance based on a fixed investment size, it is necessary to show the results on a log scale to be realistic; otherwise, the investment must change each day based on returns. Similarly, volatility looks more uniform over time when shown on a log scale.

The use of regression analysis for forecasting the price of soybeans alone had a different conclusion. Although 27 years were used, the last 10 showed a noticeable increase in soybean prices. This rising pattern is best fitted by the curvilinear and exponential models. However, these curves show price forecasts that continue to rise at an increasing rate. Had inflation maintained its double-digit rate, these forecasts would still lead
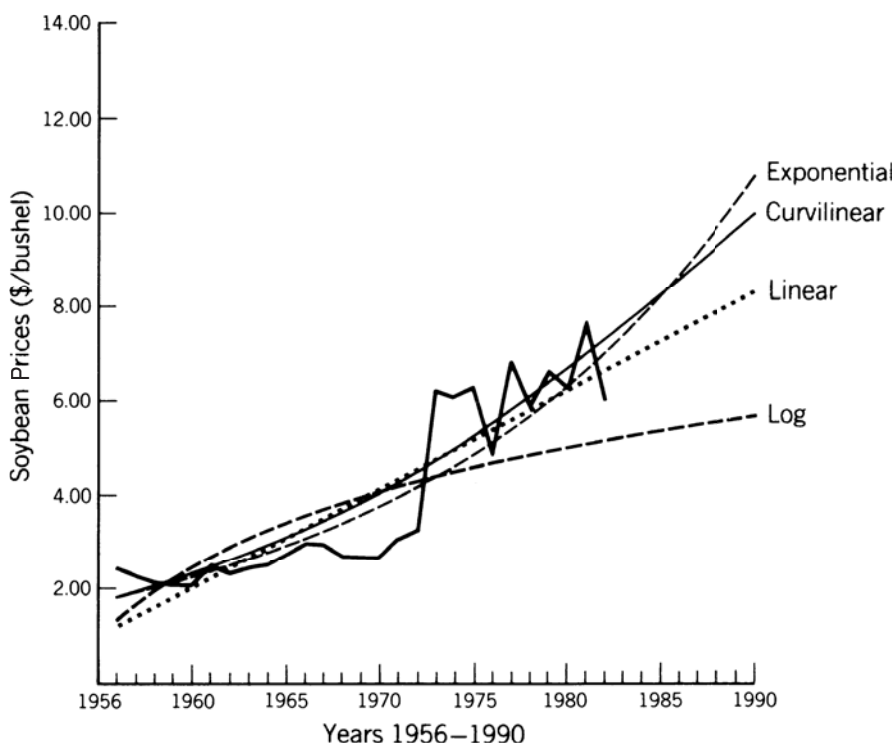
FIGURE 6.11 Least-squares approximation for soybeans using linear, curvilinear, logarithmic, and exponential models.

to unrealistically high prices. The logarithmic model, which tends to flatten at the top, turned out to be the most realistic. This shows that the problem of forecasting is more complex than this naive solution. Common sense tells us that prices cannot continue to accelerate upwards. The logarithmic model, which showed the worst statistical results, provided the best forecast. The logarithmic model is similar to the square root technique that is commonly used to estimate volatility of longer time intervals using shorter interval data, as was shown in the section of Chapter 2 called "Risk and Volatility."

To avoid using the wrong solution, most analysts carefully track the major economic factors, such as interest rates, inflation, and the value of the U.S. dollar. The model must be reestimated whenever these factors change or whenever there is new data that may result in a more accurate solution. They may use a *rolling calculation interval*, dropping off the oldest data and adding the most recent. They may also use *multivariate approximations*, discussed in the next section, which evaluate more than one independent, or causal, variables.

Direct Relationships

The price at which a commodity or stock trades is often dependent upon the prices of other competitive or substitute products. This can be seen for stocks where two

pharmaceutical companies provide the same product (such as the cholesterol drugs Lipitor and Zocor), and for airlines vying for passengers on the same route.

The ability to substitute one product for another creates an opportunity for *arbitrage*. In general, if two physical products provide the same function, they should sell at the same price, net of transaction costs, which can include carrying charges, shipping, inspection, and commissions. These relationships are watched carefully when they move apart by more than the transaction costs; traders step in quickly to buy the cheaper one and sell short the more expensive one, causing prices to come back into alignment. In the equities markets this is the basis for pairs trading. In commodities, arbitrageurs keep the futures and cash prices together; they continually prevent the price of gold in New York, London, and Hong Kong from drifting apart. For interest rate markets, "strips" serve the purpose of preventing the large pool of 3-month rates, 5- and 10-year notes, and 30-year bonds from offering widely different returns when they revert to the same maturity. The Interbank market provides the same stability for foreign exchange markets, and the soybean *crush*, energy *crack*, and other processing margins do not stay out of line for long. Competition will keep the prices of Lipitor and Zocor together, and the addition of a generic substitute will cause all three to realign. The following nonfinancial markets have close relationships that can be found using regression analysis.

Commodity Product Relationships	Reason
All crops	Income for farmers based on per acre yields
Grains	Protein value used for feed
Livestock and feedgrains	Cost of feed affects the cost of livestock
All livestock	Consumer purchases influenced by price
Sugar and corn	Commercial sweetener substitution
Hogs and pork bellies	Product dependency: bacon prices depend on hog prices
Silver, gold, and platinum	Investor's inflation hedge
Interest rates and stock markets	Investors continually choose between stocks and interest rates
Interest rates and foreign exchange	Investors worldwide move money to seek the best returns and protect against inflation

MULTIVARIATE APPROXIMATIONS

Regression analysis is most often used in complex economic models to find the combination of two or more independent variables that best explain or forecast prices. A simple application of annual production and distribution of soybeans will determine whether these factors are significant in determining the price of soybeans. Because the demand for soybeans and its products is complex, we should not expect a very accurate model using only two independent variables. However, the method of solution is the same when you add other factors. The first example will use a small sample of data, shown

in Table 6.8; at the end of this section the same method will be applied to wheat using a much longer set of data.

Applying the method of least squares which was used for a simple linear regression, the equation for two independent variables is

$$y = a_0 + a_1 x_1 + a_2 x_2$$

where y = the resulting price, in this case soybeans

x_1 = the total production (supply)

x_2 = the total distribution (demand)

a_0, a_1, a_2 = constants, or weighting factors, to be calculated

As in the linear approximation, the solution to this problem will be found by minimizing the sum of the squares of the errors at each point, where $\hat{y}_i$ is the approximation for the ith data item, and y_i is the actual price.

$$S = \sum_{i=1}^{N} \left(y_i - \hat{y}_i \right)^2$$

Substituting $\hat{y}$ from the previous equation,

$$S = \sum_{i=1}^{N} \left(y_i - \left(a_0 + a_1 x_1 + a_2 x_2 \right) \right)^2$$

The solution to the multivariate problem of two independent variables x_1 and x_2 requires the following three least-squares equations:

$$a_0 N + a_1 \sum x_1 + a_2 \sum x_2 = \sum y$$

$$a_0 \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1 x_2 = \sum x_1 y$$

$$a_0 \sum x_2 + a_1 \sum x_1 x_2 + a_2 \sum x_2^2 = \sum x_2 y$$

The procedure for solving the three simultaneous equations is the same as the curvilinear method of coefficient elimination. The sums are calculated in Table 6.8, then substituted into the last three equations:

$$12.00 a_0 + 13.546 a_1 + 12.58 a_2 = 44.4$$

$$13.55 a_0 + 15.980 a_1 + 14.75 a_2 = 53.28$$

$$12.58 a_0 + 14.756 a_1 + 13.86 a_2 = 49.264$$

TABLE 6.8 Totals for Multivariate Solution

		y	Supply x_1	Demand x_2	x_1^2	x_2^2	$x_1 x_2$	$x_1 y$	$x_2 y$
			(Billions)		(Billions)			(Millions)	
1964	1	2.67	0.700	0.677	0.490	0.458	0.474	1.869	1.808
1965	2	2.88	0.845	0.738	0.714	0.545	0.624	2.434	2.125
1966	3	2.98	0.928	0.839	0.861	0.704	0.779	2.765	2.500
1967	4	2.93	0.976	0.874	0.953	0.764	0.853	2.860	2.561
1968	5	2.69	1.107	0.900	1.225	0.810	0.996	2.978	2.421
1969	6	2.63	1.133	0.946	1.284	0.895	1.072	2.980	2.488
1970	7	2.63	1.127	1.230	1.270	1.513	1.386	2.964	3.235
1971	8	3.08	1.176	1.258	1.383	1.583	1.479	3.622	3.875
1972	9	3.24	1.271	1.202	1.615	1.445	1.528	4.118	3.894
1973	10	6.22	1.547	1.283	2.393	1.646	1.985	9.622	7.980
1974	11	6.12	1.215	1.435	1.476	2.059	1.744	7.436	8.782
1975	12	6.33	1.521	1.200	2.313	1.440	1.825	9.628	7.596
Σ		44.40	13.546	12.582	15.977	13.862	14.745	53.276	49.265

The coefficient matrix solution is[4]

$$
\begin{pmatrix}
12.00 & 13.55 & 12.58 & 44.40 \\
13.55 & 15.98 & 14.74 & 53.28 \\
12.58 & 14.74 & 13.86 & 49.26
\end{pmatrix}
$$

$$
\begin{pmatrix}
1 & 1.239 & 1.048 & 3.700 \\
0 & .6798 & .5451 & 3.145 \\
0 & .5451 & .6720 & 2.714
\end{pmatrix}
$$

$$
\begin{pmatrix}
1 & 0 & .1429 & -1.5240 \\
0 & 1 & .8018 & 4.6264 \\
0 & 0 & .2349 & .1922
\end{pmatrix}
$$

$$
\begin{pmatrix}
1 & 0 & 0 & -1.641 \\
0 & 1 & 0 & 3.9703 \\
0 & 0 & 1 & .8183
\end{pmatrix}
$$

---

[4] Matrix elimination is the necessary solution to the multivariate problem. Appendix 2 contains the computer programs necessary to perform its operation.

The results show $a_0 = -1.641$, $a_1 = 3.9703$, and $a_2 = 0.8183$, so that the multiple linear approximation of the price is

$$\hat{y} = -1.641 + 3.9703 x_1 + 0.8183 x_2$$

Where x_1 = the production of soybeans in billions of bushels
      x_2 = demand in billions of bushels

The coefficient of supply is much larger than the coefficient of demand; therefore it is the principal factor in the determination of price. Had either coefficient of x_1 or x_2 been small, it would have indicated a lack of significance. The selection of which data to use when determining price components is not always obvious and may result in a number of small coefficients. In this example, supply-and-demand figures were used to determine price, but perhaps supply and inflation or demand and inflation would have been better. Further, this method will always give an answer, even when the choice of data is erroneous. To find out which sets of data are best, each combination would have to be tested and the variance of the estimated values compared with the actual prices. The best fit and the best choice should be the one with the smallest variance, provided you have used sensible inputs.

As with the least-squares and polynomial solutions, user-friendly software as well as fully programmable platforms are available. This makes solutions available to everyone. If we look at a longer set of data for wheat, available from the Commodity Research Bureau Yearbook (published annually by John Wiley & Sons), we can take the supply as *World Production* and calculate the demand by subtracting the *Ending Stocks* from the *Production*, giving us *Usage*. Finally, the cash price in Chicago is selected as the value to be forecast. The data was available from 1971 to 2000, but only the first few years are shown in Table 6.9. The full spreadsheet and the forecasted results are available on the Companion Website as *Wheat supply and demand*.

TABLE 6.9   The First Few Years of Wheat Data Used for the Multivariate Solution, along with the Forecasted Price

Crop Year	World Production	Ending Stocks	Usage	Chicago Cash	Forecast
1970–1971	306.5	80.5	226	1.33	2.42
1971–1972	344.1	89.2	254.9	1.34	2.71
1972–1973	337.5	74.9	262.6	1.73	3.02
1973–1974	361.3	82.4	278.9	3.95	3.15
1974–1975	355.2	81.4	273.8	4.09	3.09
1975–1976	352.6	86.7	265.9	3.55	2.91
1976–1977	414.3	127.4	286.9	2.73	2.64
1977–1978	377.8	109.2	268.6	2.33	2.63
1978–1979	438.9	134.8	304.1	2.97	2.78
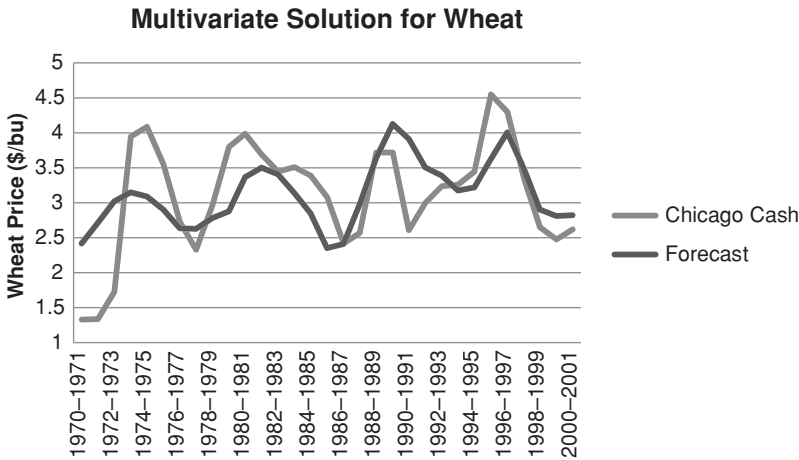
FIGURE 6.12 Multivariate solution for wheat.

The solution to the equation *Cash price$_i$* = a_0 + $a_1 \times$ *production* + $a_2 \times$ *usage* gives $a_0 = 0.290926$, $a_1 = -0.01399$, and $a_2 = 0.028385$. A chart of the forecast values compared to the actual cash prices is shown in Figure 6.12. The forecast is reasonably close to the cash prices, moving in the same direction most years but with generally lower volatility. Perhaps a better estimate of world demand, rather than relying on final U.S. stocks, would make this more accurate, then an estimate of supply and demand would result in a reasonably good estimate of the price.

Recently, *neural networks* and *genetic algorithms* have been used to find solutions to complex problems of supply and demand. Both have the advantage of providing a nonlinear solution. That is, the multivariate answer to the soybean problem says that the effect of supply is always four times greater than the effect of demand; but the neural net solution may qualify that answer by giving a larger weight to supply during years when supply is threatened by bad weather. Both genetic algorithms and neural nets are discussed in Chapter 20.

Selecting Data for an S&P Model

The stock market is driven by a healthy economy and lower interest rates. A good economy means that there is high employment and consumers are actively buying homes, durable goods, services, and frivolous items with their disposable income. Low interest rates increase corporate profitability by reducing the cost of debt and provide lower mortgages for homeowners. Economic growth and controlled inflation are delicately orchestrated by the central bank of each country; in the United States, it is the Federal Reserve.

To create a robust S&P model, one that explains the change in S&P prices using economic data, it is necessary to select the most meaningful data. The following was suggested by Lincoln[5] to be used for a 6-month S&P forecast:[6]

- *S&P prices*, the closing prices of the Standard & Poor's S&P 500 cash index.
- *Corporate bonds/Treasury bonds*, the BAA corporate bond yield divided by the 30-year Treasury bond yield, normalized by subtracting the historical mean.
- *Annual change in the U.S. dollar*, the 12-month change in the dollar, minus 1, which might be based on the Dollar Index traded on the New York Futures Exchange (discussed in Chapter 2), or a weighting of major currencies.
- *Annual change in federal funds rate*, the 12-month change in the fed funds rate, minus 1.
- *Federal funds rate/Discount rate*, the Fed funds rate divided by the discount rate, normalized by subtracting the historical mean.
- *Money supply*, M1 money supply, not seasonally adjusted.
- *Annual Consumer Price Index*, the 12-month change in the CPI, minus 1.
- *Inflation/Disinflation index*, the annualized 1-month change in the CPI divided by the 12-month change in the CPI, minus 1.
- *Leading economic indicators*, the 12-month change in the leading economic indicators, minus 1.
- *One-month versus 10-month oscillator for the S&P cash index*, the difference between the monthly average and the past 10 months, approximately 200 days.
- *Inflation-adjusted commercial loans*, the inflation-adjusted 12-month growth in commercial loans.

Lincoln used 20 years of monthly values to forecast the S&P price six months ahead. Some of this data is available on a weekly basis and might be adapted to a shorter time frame; however, it is unreasonable to think that an accurate daily forecast is possible using weekly or monthly data. It is also likely that the use of more frequent S&P price data is inconsistent with the monthly statistics and will introduce more noise and make the results less reliable.

One element that is missing from the 11 items listed is a volatility adjustment. There has been a 600% increase in the S&P price over 20 years and volatility is clearly much higher. A simple percentage relationship may not describe this change adequately, but may be used until a better one is found. Therefore, those items that become more or less volatile as prices move higher and lower must be corrected using a volatility-normalizing factor. For example, if we consider the initial volatility when the S&P was 100 at the beginning of the data as normal, then when the price is 1400 we divide the current volatility by

---

[5] Thomas H. Lincoln, "Time Series Forecasting: ARMAX," *Technical Analysis of Stocks & Commodities* (September 1991).

[6] Note that the "minus 1" in the data list refers to the calculation of returns, final value divided by starting value, minus 1.

14 to normalize, or use some volatility stabilization adjustment. This type of adjustment would apply to nearly all the items except money supply and the inflation/disinflation index. Remember that, where applicable, yields must always be used for interest rates, not prices.

Generalized Multivariate Solution

In general, the relationship between n independent variables is expressed as

$$y = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

The solution to this equation is the natural extension of the problem in two and three variables. The $n + 1$ equations in $n + 1$ variables are created by summing the $n + 1$ equations developed from the general equation by multiplying the second equation by x_1, the third by x_2, and so on:

$$a_0 n + a_1 \sum x_1 + a_2 \sum x_2 + \cdots + a_n \sum x_n = \sum y$$

$$a_0 \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1 x_2 + \cdots + a_n \sum x_1 x_n = \sum x_1 y$$

$$a_0 \sum x_2 + a_1 \sum x_1 x_2 + a_2 \sum x_2^2 + \cdots + a_n \sum x_2 x_n = \sum x_2 y$$

$$\vdots$$

$$a_0 \sum x_n + a_1 \sum x_1 x_n + a_2 \sum x_2 x_n + \cdots + a_n \sum x_n^2 = \sum x_n y$$

It is not practical to solve this system of equations manually, but it can be done quickly using a single program found in a math software library such as IMSL, Pro-Stat, MatLab, or even using Excel matrix functions (see a more detailed example in Appendix 2). Those with only a little experience in regression analysis should remember that the model is most accurate within the range of the data points; when projecting outside the bounds of the sample data, the predictive qualities of the regression formula decrease with time. It is most likely that a good solution will be found for interest rates and grain prices, which are trading at the same levels as previous years, but may not be reliable for index markets or gold that have made erratic new highs over the years.

Many dependent variables (x_i) may be used to increase the possibility of finding a good fit. The predictive quality of this solution will depend on the relevance of the independent variables. It is best to start with the obvious components of a time series, such as inflation, then add standard economic statistics, including the Consumer Price Index and Industrial Product, as well as supply and demand information specific to the market being evaluated. For both grain and energy markets, the accumulation of inventory, or stocks, is a strong influence on price; these factors also have an expected

seasonal variation, which is represented in terms of an index of adjustment. Using monthly data means that you must have a strong tolerance for risk because even the best models do not account for the price fluctuations during the month. Measuring the error of the estimates will help determine whether additional factors are necessary. When done, look at the values of a. Those that are very small should be discarded. The fewer inputs the better.

Least-Squares Sinusoidal

When you know that there is a seasonal or cyclic pattern to price movement, you can use the trigonometric functions, *sine* and *cosine*, as a special case of multiple linear equations. Observing periodic peaks and valleys in a price series suggests that a cyclic pattern may be present. One of the more well-known uses of cyclic analysis was performed by Hurst in *The Profit Magic of Stock Transaction Timing*, in which there is an interesting example of Fourier analysis applied to the Dow Jones Industrial Averages. A full discussion of cyclic analysis and trigonometric estimations can be found in Chapter 11.

The equation for the approximation of a periodic movement is

$$y_t = a_0 + a_1 t + a_2 \cos\frac{2\pi t}{P} + a_3 \sin\frac{2\pi t}{P} + a_4 t \cos\frac{2\pi t}{P} + a_5 t \sin\frac{2\pi t}{P}$$

which is a special case of the generalized multivariate approximation

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5$$

where P = the number of data points in each cycle

$x_1 = t$, the incremental time element

$x_2 = \cos(2\pi t/P)$, a cyclic element

$x_3 = \sin(2\pi t/P)$, a cyclic element

$x_4 = t\cos(2\pi t/P)$, an amplitude-variation element

$x_5 = t\sin(2\pi t/P)$, an amplitude-variation element

The number of data points, P, in each cycle would be 12 if you were calculating seasonality using monthly data. If you believe that there is a cycle, then you can find the average number of data points between major price peaks. The term $a_1 t$ will allow for the linear tendencies of the sequence. The term 2π refers to an entire cycle and $2\pi t/P$ is a section $(1/P)$ of a specific cycle t; this in turn adds weight to either the sine or cosine functions at different points within a cycle.

The solution is calculated in the tabular manner of the other methods, using simultaneous linear equations derived in the same way as the generalized multivariate equation, substituting from the table of sums and solving the coefficient matrix for $a_0, \ldots, a_5$ using the techniques in Appendix 2.

ARIMA

An *Autoregressive Integrated Moving Average* (*ARIMA*) model is created by a process of repeated regression analysis over a moving time window, resulting in a forecast value based on the new fit. An ARIMA process automatically applies the most important features of regression analysis in a preset order and continues to reanalyze results until an optimum set of parameters or coefficients is found. In Chapter 21, the selection and testing of individual parameters within a trading strategy are discussed. This involves approximating their initial value and identifying a testing range. An ARIMA model does all of this as part of its special process. Because it is used to recalculate the *best fit* each time a new piece of data appears, it may be thought of as an *adaptive process*.

G. E. P. Box and G. M. Jenkins refined ARIMA at the University of Wisconsin,[7] and their procedures for solution have become the industry standard. This technique is often referred to as the *Box-Jenkins forecast*. The two important terms in ARIMA are *autoregression* and *moving average*. *Autoregression* refers to the use of the same data to self-predict, for example, using only gold prices to arrive at a gold price forecast. *Moving average* refers to the normal concept of smoothing price fluctuations, using a rolling average of the past n days. The moving average and popular variations are discussed thoroughly in Chapters 7 and 8. This process uses an exponential smoothing technique, which is convenient for computation.

In the ARIMA process, the autocorrelation is used to determine to what extent past prices will forecast future prices. In a *first-order autocorrelation*, only the prices on the previous day are used to determine the forecast. This would be expressed as

$$p_t = a \times p_{t-1} + e$$

where p_t = the price being forecast (dependent variable)
  p_{t-1} = the price being used to forecast (independent variable)
  a = the coefficient (constant factor)
  e = the forecast error

In a *second-order autoregression*, the previous two prices are used,

$$p_t = a_1 \times p_{t-1} + a_2 \times p_{t-2} + e$$

where the current forecast p_t is based on the two previous prices p_{t-1} and p_{t-2}; there are two unique coefficients and a forecast error. The moving average is used to correct for

[7] G.E.P. Box and G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, 2nd ed. (San Francisco: Holden-Day, 1976).

the forecast error, *e*. There is also the choice of a first- or second-order moving average process,

$$\textit{First-order:} \quad E_t = e_t - be_{t-1}$$

$$\textit{Second-order:} \quad E_t = e_t - b_1 e_{t-1} - b_2 e_{t-2}$$

where E_t = the approximated error term
                           e_t = today's forecast error
            e_{t-1} and e_{t-2} = the two previous forecast errors
              b_1 and b_2 = the two regression coefficients

Because the two constant coefficients, b_1 and b_2, can be considered percentages, the moving average process is similar to exponential smoothing.

The success of the ARIMA model is determined by two factors: high correlation in the autoregression and low variance in the final forecast errors. The determination of whether to use a first- or second-order autoregression is based on a comparison of the *correlation coefficients* of the two regressions. If there is little improvement using the second-order process, it is not used. The final forecast is constructed by adding the moving average term, which approximates the errors, back into the autoregressive process

$$p'_t = p_t + E_t + e'$$

where p'_t = the new forecast
            e' = the new forecast error

The moving average process is again repeated for the new errors e', added back into the forecast to get a new value p'' and another error e''. When the variance of the errors becomes sufficiently small, the ARIMA process is complete.

The contribution of Box and Jenkins was to stress the simplicity of the solution. They determined that the autoregression and moving average steps could be limited to first- or second-order processes. To do this, it was first necessary to *detrend* the data, thereby making it *stationary*. Detrending can be accomplished most easily by *differencing* the data, creating a new series by subtracting each previous term p_{t-1} from the next p_t. Of course, the ARIMA program must remember all of these changes, or *transformations*, in order to restore the final forecast to the proper price notation by applying these operations in reverse. If a satisfactory solution is not found in the Box-Jenkins process, it is usually because the data are still not stationary and further differencing is necessary.

With the three features just discussed, the Box-Jenkins forecast is usually shown as *ARIMA* (*p, d, q*), where *p* is the number of autoregressive terms, *d* is the number of differences, and *q* is the number of moving average terms. The expression *ARIMA* (0,1,1)

is equivalent to simple exponential smoothing, a common technique discussed in the next chapter. In its normal form, the Box-Jenkins ARIMA process performs the following steps:

1. **Specification.** Preliminary steps for determining the order of the autoregression and moving average to be used:
   - *The variance must be stabilized.* In many price series, increased volatility is directly related to increased price. In stocks, the common assumption is that this relationship is *log-normal*, that increasing volatility takes the shape of a logarithmic curve. A simple test for variance stability, using the log function, is checked before more complex transformations are used.
   - *Prices are detrended.* This uses the technique of first differences; however, a second difference (or more) will be performed if it helps to remove further trending properties in the series (this is determined by later steps).
   - *Specify the order of the autoregressive and moving average components.* This fixes the number of prior terms to be used in these approximations (not necessarily the same number). In the Box-Jenkins approach, these numbers should be as small as possible; often one value is used for both. Large numbers require a rapidly expanding amount of calculation, even for a computer.

     The number of terms is a critical part of the ARIMA solution. The object of this last step is to find the fewest terms necessary to solve the problem. All ARIMA programs will print a *correlogram*, a display of the autocorrelation coefficients. The correlogram is used to find whether all the trends and well-defined periodic movements have been removed from the series by differencing. Figure 6.13 shows the output from ProStat for 15 lags using the
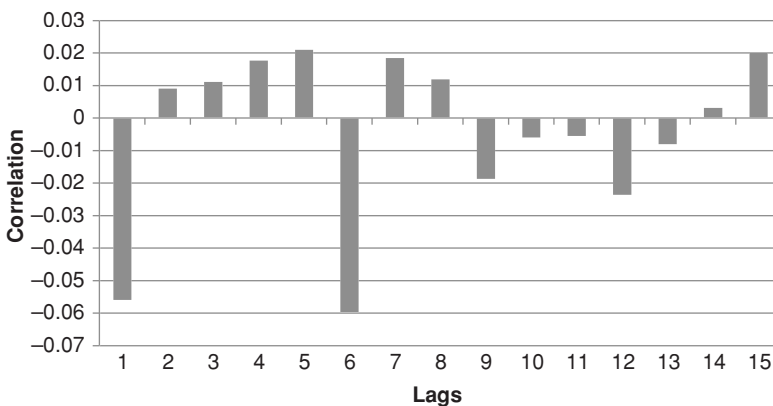
FIGURE 6.13   ARIMA correlograms from autocorrelations. Lags shown along the bottom give correlations at the left. None of them are significant (output from ProStat based on gold futures).

first differences of 20 years of gold futures prices. If any of the correlations had been significant (for example, over 0.40), the data would be differenced again. The correlogram is one way of visualizing serial correlation, the dependence of current data on previous data. This same objective will be accomplished differently using a Monte Carlo process to estimate investment risk, discussed in Chapter 23.

2. **Estimation: determining the coefficients.** The previous step was used to reduce the number of autoregressive and moving average terms necessary to the estimation process. The ARIMA method of solution is one of minimizing the errors in the forecast. In minimization, it will perform a linear or nonlinear regression on price (depending on the number of coefficients selected), determine the errors in the estimation, and then approximate those errors using a moving average. It will next look at the resulting new error series, attempt to estimate and correct the errors in that one, and repeat the process until it cannot improve results further.

To determine when an ARIMA process is completed, three tests are performed at the end of each estimation pass:

1. *Compare the change in the coefficient value.* If the last estimation has caused little or no change in the value of the coefficient(s), the model has successfully converged to a solution.

2. *Compare the sum of the squares of the error.* If the error value is small or if it stays relatively unchanged, the process is completed.

3. *Perform a set number of estimations.* Unless a maximum number of estimations is set, an ARIMA process might continue indefinitely. This safety check is necessary in the event the model is not converging to a solution.

Once completed, the errors can be examined using an *O*-statistic to check for any trend. If one exists, an additional moving average term may be used to eliminate it.

Forecast Results

Once the coefficients have been determined, they are used to calculate the forecast value. These forecasts are most accurate for the next day and should be used with less confidence for subsequent days (see Figure 6.14).

What if the forecast does not work? First, the process should be checked to be certain that it was performed properly. Be sure that any data transformations were reversed. Pay particular attention to the removal of trends using the correlogram. Next, check the data used in the process. If the data sample is changing (which can be observed on a price chart), select either a shorter or longer period that contains more homogeneous data, that is, data similar to the current market period.
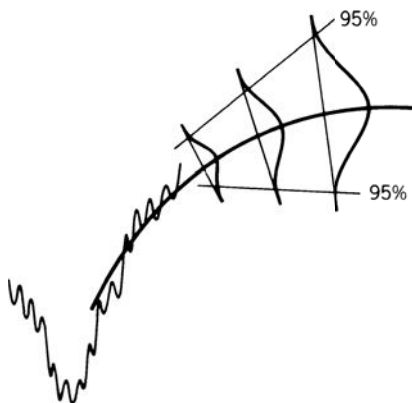
FIGURE 6.14   ARIMA forecast becomes less accurate as it is used further ahead.

ARIMA Trading Strategies

In the article that originally piqued the interest of traders,[8] Anon uses a 5-day-ahead forecast. If the ARIMA process forecasts an uptrend and prices fall below the forecast value, the market can be bought with added confidence (expecting lower risk and more profit by buying at a price that is below estimated value). This technique of selecting better entry points may compensate for some of the inaccuracies latent in any forecasting method. The danger of this approach is that prices may continue to move counter to the forecast.

Following the Trend
Use the 1-day-ahead forecast to determine the trend position. Hold a long position if the forecast is for higher prices, and take a short position if the process is expecting lower prices.

Mean-Reverting Indicator
Use the ARIMA confidence bands to determine overbought/oversold levels. Not only can a long position be entered when prices penetrate the *lowest* 95% confidence band, but they can be closed out when they return to the normal 50% level. Although mean reversion trades are tempting, they always carry more risk than trend trading. A conservative trader will enter the market only in the direction of the ARIMA trend forecast. As shown in Figure 6.14, if the trend is up, only the penetrations of a lower confidence band will be used to enter new long positions.

---

[8] Louis J. Anon, "Catch Short-Term Profits with ARIMA," *Commodities Magazine* (December 1981).

Use of Highs and Lows

The implied highs and lows, as well as the independently forecasted highs and lows, can be the basis for other interesting strategies.[9] The following two are used with intraday prices.

1. Using confidence bands based on the closing prices, buy an intraday penetration of the expected high or sell short a penetration of the expected low, and liquidate the position on the close. Use a stop-loss. Take positions only in the direction of the ARIMA trend.

2. Using the separate ARIMA models based on the daily high and low prices, buy a penetration of the 50% level of the high and sell a penetration of the 50% level of the lows. Liquidate positions on the close. Use a stop-loss.

Slope

The 1-day-ahead forecast suggested in "Following the Trend" a few paragraphs earlier is essentially a projection of the slope of the trendline. The purpose of directional analysis, whether regression or moving averages, is to uncover the true direction of prices by discarding the noise. Therefore, the slope of the trendline, or the direction of the regression forecast, is the logical answer.

The popular alternate for triggering a new directional signal is a price penetration of an envelope or band value. Using regression analysis, that band can be replaced by a confidence level. While it is true that the number of random, or false, penetrations declines as the confidence band gets farther away from the trendline, so does the total number of penetrations. At any band distance, there are still a large number of erroneous signals. The slope itself should be viewed as the best approximation of direction.

Kalman Filters

Kalman offers an alternative approach to ARIMA, allowing an underlying forecasting model (*message model*) to be combined with other timely information (*observation model*). The message model may be any trading strategy, moving average, or regression approach. The observation model may be the specialist's or floor broker's opening calls, market liquidity, or earlier trading activity in the same stock or index market on a foreign exchange—all of which have been determined to good candidate inputs for forecasting.

Assume that the original forecast (message) model can be described as

$$M\left(p_t\right) = c_f\, p_{t-1} + me_t$$

and the observation model as

$$O\left(p_t\right) = c_0\, p_t + oe_t$$

where *me* and *oe* are the message and observation model errors, respectively

---

[9] John F. Kepka, "Trading with ARIMA Forecasts," *Technical Analysis of Stocks & Commodities* (August 1985).

The combined forecast would then use the observation model error to modify the result

$$p'_{t+1} = c_f\, p'_t + K_{t+1} o e_t$$

where K is the Kalman gain coefficient,[10] a factor that adjusts the error term.

BASIC TRADING SIGNALS USING A LINEAR REGRESSION MODEL

A linear regression, or straight-line fit, could be the basis for a simple trading strategy similar to a moving average. For example, an n-day linear regression, applied to the closing prices, can produce a 1-day-ahead forecast price, F_{t+1}, a projection of the slope. We can use this with the following rules for trading:

- *Buy* when tomorrow's closing price (C_{t+1}) moves above the forecasted value of tomorrow's close (F_{t+1}).
- *Sell short* when tomorrow's closing price (C_{t+1}) moves below the forecasted value of tomorrow's close (F_{t+1}).[11]

Adding Confidence Bands

Because the linear regression line passes through the center of price movement during a steady period of rising or falling prices, these rules would produce a lot of buy and sell signals. To reduce the frequency of signals and avoid changes of direction due to market noise, confidence bands are drawn on either side of the regression line. A 90% confidence band is simply 1.65 times the standard deviation of the residuals (R_t, the difference between the actual prices and the corresponding value on the regression line as of time t, the most recent price). A 95% confidence band uses a multiplier of 1.96; however, most people use 2.0 simply for convenience. The new trading rules using a 95% confidence band would then become

- *Buy* when tomorrow's closing price (C_{t+1}) moves above the forecasted value of tomorrow's close (F_{t+1}) + ($2.0 \times R_t$).
- *Sell short* when tomorrow's closing price (C_{t+1}) moves below the forecasted value of tomorrow's close (F_{t+1}) − ($2.0 \times R_t$).

An important difference between a model based on linear regression and one founded on a moving average is the lag. Both methods assume that prices will continue to move according to the pattern identified using the last price data available. If prices continue

---

[10] For a more complete discussion, see Andrew D. Seidel and Philip D. Ginsberg, *Commodities Trading* (Englewood Cliffs, NJ: Prentice-Hall, 1983), or R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering* (March 1960).

[11] These rules were used by Frank Hochheimer and Richard Vaughn in *Computerized Trading Techniques 1982* (New York: Merrill Lynch Commodities, 1982).
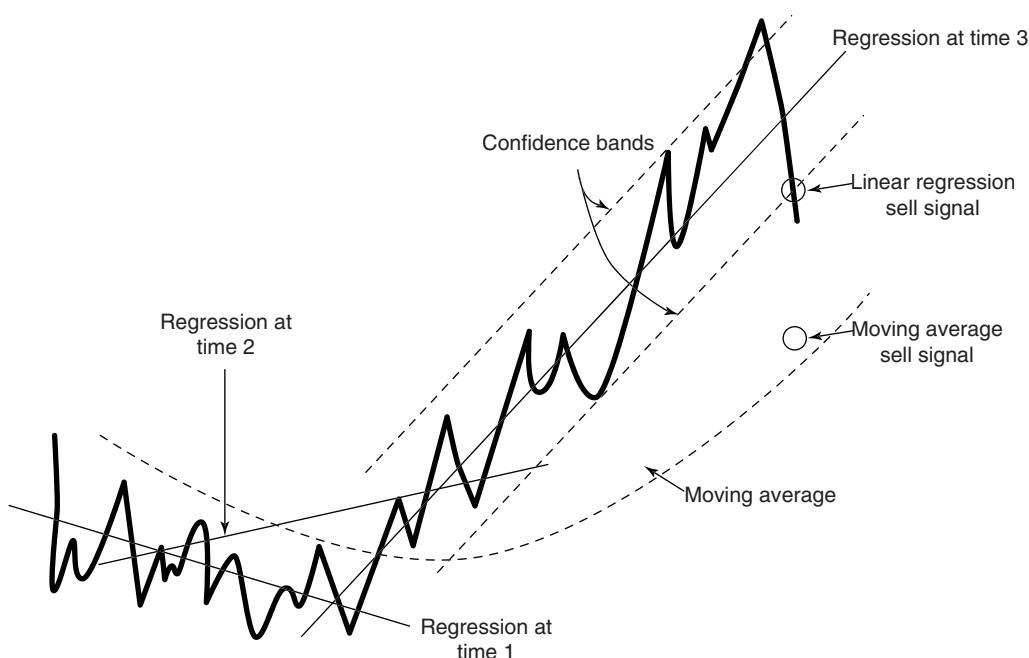
FIGURE 6.15 Linear regression model. Penetration of the confidence band turns the trend from up to down. When prices move steadily up, the regression model will signal a change of direction faster than a moving average.

higher at the same rate, a moving average system will initially lag behind, then increase at the same rate. The lag creates a safety zone to absorb some minor changes in the direction of prices without indicating that the trend has changed. (See Chapter 7 for a complete discussion of moving averages, and Chapter 8 for a comparison of a linear regression slope trading system with five other popular trending methods.)

A regression model, however, identifies a change of direction sooner by measuring tomorrow's actual price against the projected future price (a straight-line projection for a linear regression). Confidence bands around the straight-line projection will decide the size of the price move up or down needed to change the trend direction. Figure 6.15 shows the changing direction of a rolling linear regression at three points in time compared to a moving average. At the most recent period on the chart the reversal point for the trend direction is much closer using the confidence bands of the regression than the lagged moving average. This may not be the case during the intervals where price changes direction at the bottom of the chart. The Companion Website has a spreadsheet, *Bund regression with bands*, that produces trading signals and performance using regression bands. It also compares the results with the slope method described in the next section.

In general, the use of regression lines works well for finding the price trend, although actual trading signals differ from moving average and breakout strategies. As with most trending systems, performance tends to improve as the calculation period increases, capturing the largest economic trends.

IBM(D) - Daily NYSE L= 162.88 -0.34 -0.21% O=0.00 Hi=0.00 Lo=0.00 V=1,400 Mov Avg 1 Line (Close, 60,0) 150.87

TSMS Slope (60) 0.40

TSMS R2(60) 0.1333 0.0956

FIGURE 6.16   IBM trend using the slope and R.

Using the Linear Regression Slope

The slope of the linear regression line, the angle at which it is rising or falling, is an effective way to simplify the usefulness of the regression process. The slope shows how quickly prices are expected to change over a unit of time. The unit of time is the same as the period of data used to find the regression values. When you use a longer calculation period, the slope of the regression line changes slowly; when you use only a few days in your calculation, the slope changes quickly. Using only the slope, you can trade with the following rules:

- *Buy* when the value of the slope is positive (rising).
- *Sell* when the value of the slope is negative (falling).

When using the slope, there is less need for a confidence band because the longer calculation periods smooth out erratic price movement. The scale of the slope is different from the prices; therefore it will be seen in a separate panel on a chart (the middle panel in Figure 6.16). The linear regression slope will be compared to other trending techniques in Chapter 8 and its performance is competitive.

Adding Correlations

The slope is a practical substitution for the regression line and confidence bands, but it can also be misleading. It is possible to have the same slope value while prices are moving higher in both a smooth pattern or in a very noisy, erratic one. Using the correlation coefficient, R^2, will provide a measurement of the consistency of the price movement.[12]

---

[12] See Rudy Tesco, "LRS + R-Squared = the 95% Solution," *Technical Analysis of Stocks & Commodities* (February 2003).

Instead of R^2, which is always positive, R will be used. And, instead of using the usual price differences (or return differences), the actual prices will be used and correlated against a series of sequential numbers. This should give us a better value for whether the prices are trending. If R becomes negative then prices are trending down. This can be seen in the bottom panel of Figure 6.16. A 60-day moving average of prices is shown at the top, a corresponding 60-day slope is in the middle, and a 60-day R is at the bottom along with a 30-day (R-*period*/2) trend of R.

The trading rules for the slope are normally the same as those for a moving average: a long position is entered when the trendline or slope turns up and a short sale begins when the trendline turns down; however, the slope today can be less than yesterday but still positive, which means that the trend is still up. Then another rule would be to buy when the slope turns positive and sell short when it turns negative.

Figure 6.16 shows that using the moving average trend is good, but has the greatest lag. The slope is much faster but less stable. Adding R gives a different view of where the trends change and can complement either method.

Forecast Oscillator

Tuschar Chande used the regression forecast and its residuals to create trend-following signals called the *Forecast Oscillator*.[13] Using a 5-day regression, find the residuals and calculate the percentage variation from the regression line. A buy signal occurs when the 3-day average of the residuals crosses above the regression line; a short sale is when the 3-day average of the residuals crosses below the regression line. If

$$\% F_t = 100 \ \times \ \frac{y_t - \hat{y}_t}{y_t}$$

and $\%F_t(3)$ is the 3-day moving average of $\%F$, then

buy when $\%F_t(3)$ crosses above $\hat{y}_t$
sell short when $\%F_t(3)$ crosses below $\hat{y}_t$

This makes the assumption that the residuals have a trending quality.

MEASURING MARKET STRENGTH

One of the natural applications of the linear regression is to measure and compare the strength of one market against another. For example, we might want to ask, "Which market is leading the other, Hewlett Packard or Dell? The slope of the linear regression,

---

[13] Tushar S. Chande and Stanley Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994).
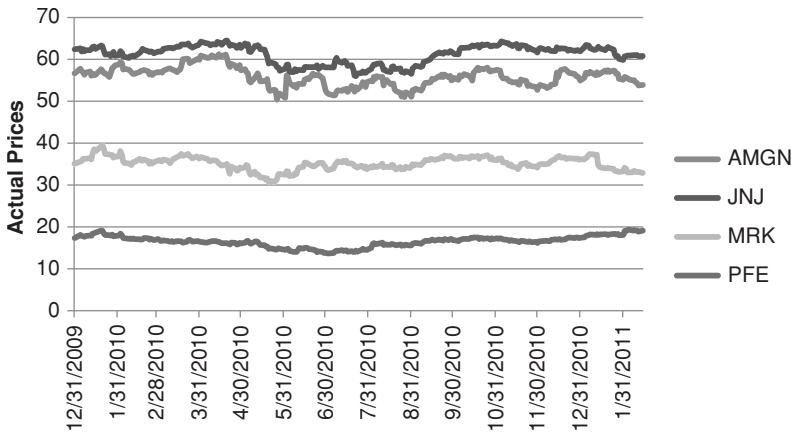
FIGURE 6.17 Prices of four pharmaceutical companies, 2010 through January 2011.

measured over the same calculation period, is the perfect tool for comparing, or *ranking*, a set of markets. Figure 6.17 shows the prices of four pharmaceutical companies, Amgen (AMGN), Johnson & Johnson (JNJ), Merck (MRK), and Pfizer (PFE). Which of them is the strongest during the last 60 days? No doubt it's PFE, the line at the bottom. But the weakest is not as clear.

In order to rank these correctly, it is best to index the stock prices, starting at 100 on the same date. For only stocks, indexing is not necessary, but if you are mixing stocks and futures, or just using futures, indexing will be needed to put the trading units into the same notation. Figure 6.18 shows the indexed stocks beginning at the value of 100.

Once indexed, apply a rolling linear regression slope to each market using the same calculation period, 60 days in this example. Figure 6.19 shows that that PFE was the most
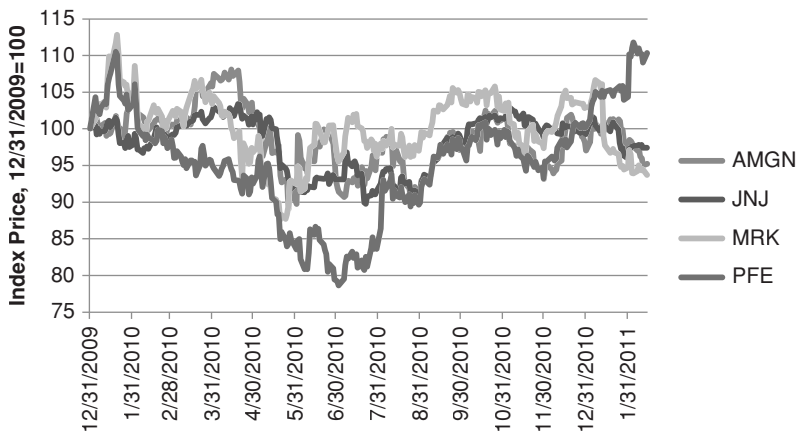
FIGURE 6.18 Indexed stocks beginning at 100 give a better view of relative performance.
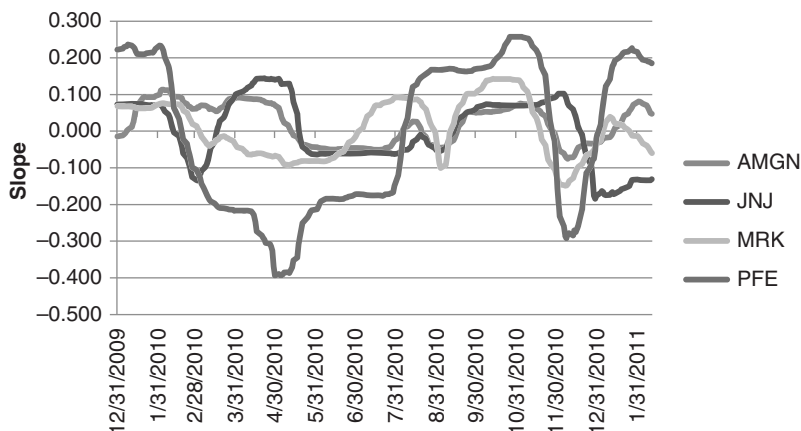
FIGURE 6.19   Rolling linear regression slope, using a 60-day calculation period, shows that each market changes from stronger to weaker over the 1-year period.

volatile because the slope varied from much higher to much lower than the others, while each stock had its turn at being the strongest and weakest. At the right side of the chart, PFE finishes the strongest followed by AMGN, MRK, and JNJ. In Chapter 13 this technique will be used to create a classic market neutral strategy. The basis of the strategy is to either buy the strongest stocks and sell short the weakest, or sell short the strongest and buy the weakest. In either case, strongest and weakest are determined by the slope of the regression rather than a moving average trend.

Time-Based
Trend Calculations

T he purpose of all trend identification methods is to remove the underlying noise in the market, those erratic moves that seem to be meaningless, and find the current direction of prices. But trends are somewhat dependent upon your time horizon. There may be more than one trend at any one time, caused by short-term events and long-term policy, and it is likely that one trader will search for the strongest, or most dominant trend, while another will seek a series of shorter-term moves. There is no "right" or "wrong" trend but a choice of benefits and compromises.

The technique that is used to uncover a particular trend can depend upon whether any of the underlying trend characteristics are known. Does the stock or futures market have a clear seasonal or cyclic component, such as the travel industry or coffee prices; or does it respond to long-term monetary policy based on the cost of servicing debt or interest income? If you know more about the reasons why prices trend, you will be able to choose the best method of finding the trend and the calculation period.

Chapter 6 used regression to find the direction of a single price series (based on one market and time), the relationship between two markets, and the ranking of both similar and diverse markets. Once you know that there is a fundamental relationship between data, a formula can be found that expresses one price movement in terms of the other prices and economic data. The predictive qualities of these methods are best when applied to data that have been seen before, that is, prices that are within the range of historic data. Forecasting reliability decreases when values are extrapolated outside the previous range of value. The most popular trend models, discussed in this chapter, ignore the reasons why trends exist and generalize the process of smoothing prices in order to find a trend's direction.

FORECASTING AND FOLLOWING

There is a clear distinction between forecasting the trend and finding the current trend. Forecasting, predicting the future price, is much more desirable but very complex. As

shown in the previous chapter, it involves combining those data that are most important to price change and assigning a value to each one. The results are always expressed with a confidence level, the level of uncertainty in the forecast. There is always lower confidence as you try to forecast further into the future.

The techniques most commonly used for evaluating the direction or tendency of prices both within prior ranges or at new levels are called *autoregressive* functions. Unlike forecasting models, they are only concerned with evaluating the current price direction. This analysis concludes that prices are moving in an upward, downward, or sideways direction, with no indication of confidence. From this simple basis, it is possible to form rules of action and develop complex trading strategies.

All of these techniques make the assumptions that past data can be used to predict future price movement, and that the direction of prices today is the most likely forecast of the direction of prices tomorrow. For the most part, these assumptions have proved to be true. From a practical viewpoint, these trending methods are more flexible than the traditional regression models, but to achieve success they introduce a lag. A *lag* is a delay in the identification of the trend. Great effort has been spent trying to reduce this lag in an attempt to identify the trend sooner; however, the lag is the zone of uncertainty that allows the technique to ignore most of the market noise.

In an autoregressive model, one or more previous prices determine the next sequential price. If p_t represents today's price, p_{t-1} yesterday's price, and so on, then tomorrow's expected price will be

$$p_{t+1} = a_0 + a_1 p_t + a_2 p_{t-1} + \ldots + a_t p_1 + e$$

where each price is given a corresponding weight a_i and then combined to give the resultant price for tomorrow $p_{t+1} + e$ (where e represents an error factor, usually ignored). The simplest example is the use of yesterday's price alone to generate tomorrow's price:

$$p_{t+1} = a_0 + a_1 p_{t-1} + e$$

which you may also recognize as the formula for a straight line, $y = a + bx$, plus an error factor.

The autoregressive model does not have to be linear; each prior day can have a nonlinear predictive quality. Then each expected price p_{t+1}, could be represented by a curvilinear expression, $p_{t+1} = a_0 + a_1 p_t + a_2 p_{t-1}^2 + e$, or by an exponential or logarithmic formula, $\ln p_{t+1} = a_0 + a_1 \ln p_t + a_2 \ln p_{t-1} + e$, which is commonly used in equity analysis. Any of these expressions could then be combined to form an autoregressive forecasting model for p_{t+1}. In going from the simple to the complex, it is natural to want to know which of these choices will perform best. Theoretically, the best method will be the one that, when used in a strategy, yields the highest return for the lowest risk; however, every investor has a personal risk preference. The answer can only be found by applying and comparing different methods, and experiencing how they perform when actually traded. It turns out that the best historic results often comes from *overfitting* the data, and is a poor choice for trading. At the end of Chapter 8 there is a comparison of popular trending systems,

and Chapter 21 will show testing methods that are most likely to lead to robust results. Throughout the book there will be comparisons of systems and methods that are similar.

Least-Squares Model

The least-squares regression model is the same technique that was used in the previous chapter to find the relationship between two markets—Barrick Gold and cash gold, corn and soybeans—or to find how price movement could be explained by the main factors influencing them, supply and demand. Most trading systems depend only on price; therefore, we will look again using the least-squares model with time as the independent variable and price as the dependent variable. The regression results will be used in an autoregressive way to forecast the price *n*-days ahead, and we will look at the accuracy of those predictions. The slope of the resulting straight line or curvilinear fit will determine the direction of the trend.

Error Analysis

A simple error analysis can be used to show how time works against the predictive qualities of regression, or any forecasting method. Using 10 years of General Electric (GE) prices, ending February 15, 2010, the slope and *y*-intercept are calculated for a rolling 20 days. The 1-, 2-, 3-, 5-, and 10-day ahead forecast is found by projecting the slope by that number of days. The forecast error is the difference between the projected price and the actual price. Figure 7.1 shows the price for GE from December 31, 2010, through February 15, 2011, along with the five forecast prices. Even though the price move seems to be increasing steadily, the forecasts get farther away as the days ahead become larger. This result is typical of forecasting error, regardless of the method, and argues that the smallest forecast interval is the best.
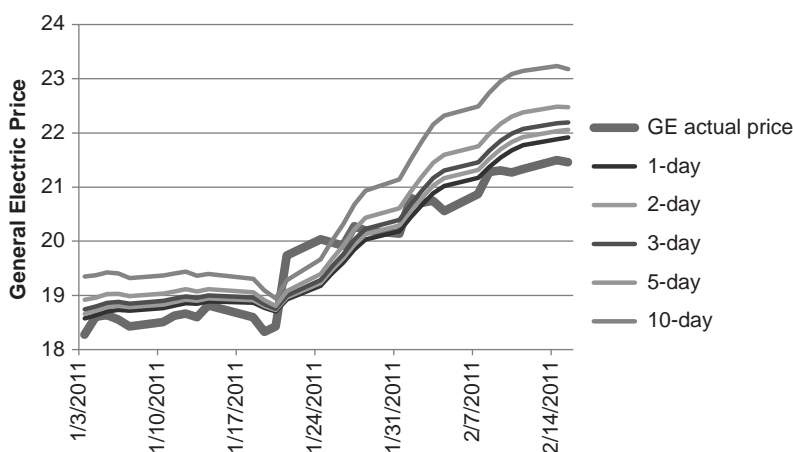
FIGURE 7.1 General Electric price from December 31, 2010, through February 15, 2011. Forecast accuracy decreases as the forecast period increases.

TABLE 7.1 General Electric Analysis of Regression Error Based on a 20-Day Rolling Calculation Period

Seq	Date	GE Close	Regression		Forecast					Forecast error				
			Slope	Intercept	1-day	2-day	3-day	5-day	10-day	1-day	2-day	3-day	5-day	10-day
n	2/1/2011	20.80	0.118	−351.689	20.438	20.556	20.675	20.911	21.502	−0.616	−0.669	−0.769	−1.128	−1.404
n+1	2/2/2011	20.71	0.131	−390.484	20.673	20.804	20.934	21.195	21.848	−0.272	−0.420	−0.478	−0.791	−1.406
n+2	2/3/2011	20.75	0.141	−422.673	20.886	21.027	21.168	21.449	22.154	−0.077	−0.194	−0.354	−0.535	−1.657
n+3	2/4/2011	20.56	0.144	−432.439	21.019	21.162	21.306	21.594	22.314	0.326	0.244	0.115	−0.128	−1.616
n+4	2/7/2011	20.87	0.146	−440.206	21.166	21.313	21.459	21.752	22.484	0.149	0.157	0.064	−0.262	−1.588
n+5	2/8/2011	21.28	0.152	−458.803	21.367	21.519	21.672	21.976	22.738	−0.114	−0.118	−0.112	−0.369	−1.615
n+6	2/9/2011	21.31	0.156	−471.628	21.544	21.700	21.857	22.169	22.952	0.057	0.003	−0.004	−0.115	−1.301
n+7	2/10/2011	21.27	0.157	−471.985	21.675	21.831	21.988	22.301	23.084	0.274	0.249	0.189	0.179	−0.957
n+8	2/11/2011	21.33	0.152	−459.262	21.769	21.921	22.074	22.379	23.141	0.345	0.370	0.342	0.264	−0.654
n+9	2/14/2011	21.50	0.150	−452.185	21.879	22.029	22.180	22.480	23.231	0.269	0.331	0.357	0.252	−0.566
n+10	2/15/2011	21.46	0.140	−419.686	21.914	22.054	22.193	22.473	23.173	0.419	0.461	0.528	0.516	−0.322

The method of finding the forecast error is shown in Table 7.1 (the full spreadsheet can be found on the Companion Website as *TSM General Electric regression error forecast*). Only the closing prices are needed, and they appear in column 3. The slope and intercept use the sequential numbers in column 1 for *X*, and the GE prices for *Y*. The *n*-day ahead forecast is

$$y_{t+n} = Intercept_t + Slope_t \times Price_t + n \times Slope_t$$

The standard deviations of the five forecast errors (Table 7.2), taken over the entire 10 years, shows the error increasing as the days-ahead increase. This confirms the expectation that forecasting accuracy decreases with time and that confidence bands will get wider with time. For this reason, any forecasts used in strategies will be 1-day ahead.

Limiting the Forecast to Direction

To be profitable trading a trending system, it is only necessary to be correct in one of the following two cases:

1. In more than 50% of the days you are correct in predicting whether prices will go up or down and the average up move is equal to the average down move.
2. Your forecast accuracy is less than 50% but the size of the correct moves is greater than the size of the incorrect moves.

Unfortunately, there is no way to prove that a particular method of forecasting, moving average, regression, or other techniques will be accurate over all calculation periods. The fact is that some calculation periods are profitable and others are not. Those that are profitable must satisfy one of the two conditions stated above. Therefore, the answer is *it's true when it works and not true when it doesn't work.*

Experience shows that the best choice of trending method will be the one that is profitable over most markets and most calculation periods. Even then, shorter calculation periods have no trend (this was discussed in Chapter 1), so we need to restrict our statement to longer periods.

The methods discussed in the remainder of this chapter are all intended to identify the direction of prices based on past data. The trading systems that use them all assume that there is a better chance of prices continuing in that direction the next day. If the systems are profitable then the assumption was correct. There is sufficient performance history for macrotrend funds that justify this conclusion. These systems will be discussed in the next chapter.

TABLE 7.2 The Standard Deviation of Errors for Different "Days Ahead" Forecasts

Days Ahead	1	2	3	5	10
Stdev of Errors	1.137	1.336	1.519	1.860	2.672

PRICE CHANGE OVER TIME

The most basic of all trend indicators is the change of price over some period of time. This is written as

$$M_t = p_t - p_{t-n}$$

where M is called *momentum*, t is today, and n is the number of days. Sometimes this is called *rate of change*, but both momentum and rate of change are incorrect names. In mathematics, this is the *first difference*, and in physics it is *speed* (distance covered over time). The next chapter will discuss this and many variations of momentum.

   If the change in price is positive, we can say that the trend is up, and if negative, the trend is down. Of course, this decision is based only on two data points, but if those points are far enough away from each other, that is, if n is large, then the trend as determined by this method will be very similar to a simple moving average, discussed in the next section. The simplest method can often be the most robust; therefore, as you read about many other approaches to analyzing price, keep asking, "Is it better than the simple change in price?"

THE MOVING AVERAGE

The most well-known of all smoothing techniques, used to remove market noise and find the direction of prices, is the *moving average (MA)*. Using this method, the number of elements to be averaged remains the same, but the time interval advances. This is also referred to as a *rolling calculation period*. Using a series of prices, $p_0, p_1, p_2, \ldots, p_t$, a moving average measured over the most recent n of these prices, or data points, at time t would be

$$MA_t = \frac{p_{t-n+1} + p_{t-n+2} + \cdots + p_t}{n} = \frac{1}{n} \sum_{i=t-n+1}^{t} p_i, \; n \leq t$$

Then today's moving average value is the average (arithmetic mean) of the most recent n data points. For example, using three points $(n = 3)$ to generate a moving average starting at the beginning of the data:

$$MA_3 = (p_1 + p_2 + p_3)/3$$
$$MA_4 = (p_2 + p_3 + p_4)/3$$
$$\vdots$$
$$MA_t = (p_{t-2} + p_{t-1} + p_t)/3$$

If p_t represents a price at a specific time t, the moving average would smooth the price changes. When more prices, n, are used in the average, each price will be a smaller part of the average and have less effect on the final value. Using five successive prices is called a *5-day moving average.* When the next sequential price is added and the oldest

is dropped off, the prior average is changed by ⅕ of the difference between the old and the new values. If

$$MA_5 = (p_1 + p_2 + p_3 + p_4 + p_5)/5$$

and

$$MA_6 = (p_2 + p_3 + p_4 + p_5 + p_6)/5$$

then $c = p_2 + p_3 + p_4 + p_5$ can be substituted for the common part of the moving average. MA_6, the most recent value, can be solved in terms of MA_5, the previous value, to get

$$MA_6 = MA_5 + (p_t - p_{t-n})/5$$

This also gives a faster way to calculate a moving average. It can be seen that the more terms in the moving average, the less effect the addition of a new term will have:

$$MA_t = MA_{t-1} + (p_t - p_{t-n})/n$$

The selection of the number of terms, called the *calculation period*, is based on both the predictive quality of the choice (measured by the error but more often by the profitability) or the need to determine price trends over specific time periods, such as a season. For outright trading, the calculation period is chosen for its accuracy in identifying trend and the risk tolerance of the trader. Slower trends, using longer calculation periods, are usually better indicators of price direction, but involve larger risk. The stock market has adopted the 200-day moving average as its benchmark for direction; however, traders find this much too slow for timing buy and sell signals.

The length of a moving average can be tailored to specific needs. A 63-day moving average, ¼ of 252 business days in the year, would reflect quarterly changes in stock price, minimizing the significance of price fluctuations within a calendar quarter. A simple yearly calculation period, 252 days, would ignore all seasonality and emphasize the annual growth of the stock. Any periodic cycle that is the same length as the moving average length is lost; therefore, if a monthly cycle has been identified, then a moving average of less than 10 days (half the cycle length) would be best for letting the moving average show that cycle. Using a moving average to find seasonal and cyclic patterns is covered in Chapters 10 and 11. At this point it is sufficient to remember that if there is a possibility of a cyclic or seasonal pattern within the data, care should be taken to select a moving average that is out of phase with that pattern (that is, not equal to the cycle period).

The length of the moving average may also relate to its commercial use. A jeweler may purchase silver each week to produce bracelets. Frequent purchases of small amounts keeps the company's cash outlay small. The purchaser can wait a few extra days during a week while prices continue to trend downward but will buy immediately when prices turn up. A 6-month trend cannot help him because it gives a long-term answer to a short-term problem; however, a 5-day moving average may give the trend direction within the jeweler's time frame.

User-Friendly Software

Fortunately, we have reached a time when it is not necessary to perform these calculations the long way. Spreadsheet programs and specialized testing software provide simple tools for performing trend calculations as well as many other more complex functions discussed in this book. The notation for many of the different spreadsheets and software is very similar and self-explanatory:

Function	Spreadsheet Notation	TradeStation Notation
Sum	=sum(list)	summation(value, period)
Moving average	=average(list)	average(value, period)
Standard deviation	=stdev(list)	stddev(value, period)
Maximum value	=max(list)	highest(value, period)
Minimum value	=min(list)	lowest(value, period)

In the spreadsheet notation, *list* is a series of rows. For example, (D11:D30) would be 20 rows in column D, and in *TradeStation* notation, *period* is the calculation period and *value* is the closing daily price series, *close* or *C*.

What Do You Average?

The closing or daily settlement is the most common price applied to a moving average. It is generally accepted as the "true" price of the day and is used by many analysts for calculation of trends. It is the price used to reconcile brokerage accounts at the end of the day, create the *Net Asset Value* (NAV) for funds, and for futures trading it is called *marked-to-market accounting*. A popular alternative is to use the average of the high, low, and closing prices, representing some sort of center of gravity. You may also try the average of the high and low prices, ignoring the closing price entirely.

Another valid component of a moving average can be other averages. For example, if p_1 through p_t are prices, and MA_t is a 3-day moving average on day t, then

$$MA_3 = (p_1 + p_2 + p_3)/3$$
$$MA_4 = (p_2 + p_3 + p_4)/3$$
$$MA_5 = (p_3 + p_4 + p_5)/3$$

and

$$MA_5' = \left(MA_3 + MA_4 + MA_5 \right)/3$$

where MA_5' is a *double-smoothed* moving average, which gives added weight to the center points. Double smoothing can be very effective and is discussed later in this chapter. Smoothing the highs and lows independently is another technique that creates a representation of the daily trading range, or volatility. This has been used to identify normal and extreme moves, and is also discussed in Chapter 8.

Types of Moving Averages

Besides varying the length of the moving average and the elements that are to be averaged, there are a great number of variations on the simple moving average. In the methods that follow, the notation assumes that the most recent day is t and the average is found over the past n days.

The Simple Moving Average

The simple moving average is the average (mean) of the most recent n days. It has also been called a *truncated moving average*, and it is the most well-known and commonly used of all the methods. Repeating the formula from earlier in this chapter,

$$MA_t = \frac{p_{t-n+1} + p_{t-n+2} + \cdots + p_t}{n} = \frac{1}{n}\sum\nolimits_{i=t-n+1}^{t} p_i, \; n \le t$$

The main objection to the simple moving average is its abrupt change in value when an important old piece of data is dropped off, especially if only a few days are used in the calculation. We also know that, if the new data, p_t, is greater than the oldest data item that will be dropped off, p_{t-n}, then the new average, MA_t will be greater than the previous average, MA_{t-1}.

Average-Modified or Average-Off Method

To avoid the end-off problem of the simple moving average, each time a new piece of data is added the previous average can be dropped off. This is called an *average-modified* or *average-off* method. It is computationally convenient because you only need to keep the old average value rather than all the data that was used to find the average. In general, the average-off method is

$$AvgOff_t = \frac{(n-1) \times AvgOff_{t-1} + p_t}{n}$$

The substitution of the moving average value for the oldest data item tends to smooth the results even more than a simple moving average and dampens the end-off impact.

Weighted Moving Average

The *weighted moving average* opens many possibilities. It allows the significance of individual or groups of data to be changed. It may restore perceived value to parts of a data sample, or it may incorrectly bias the data. A weighted moving average is expressed in its general form as

$$W_t = \frac{w_1 P_{t-n+1} + w_2 P_{t-n+2} + \ldots + w_{n-1} P_{t-1} + w_n P_t}{w_1 + w_2 + \ldots + w_n} = \frac{\sum\limits_{i=1}^{n} w_i P_{t-i+1}}{\sum\limits_{i=1}^{n} w_i}$$

The weighted moving average at time t is the average of the previous n prices, each price having its own weighting factor w_i. There is no restriction on the values used as

FIGURE 7.2 A comparison of moving averages. The simple moving average, linearly weighted average, triangular weighted, and average off methods are applied to the S&P, April through December 2010.

weighting factors; that is, they do not have to be percentages that all total to 1. The most popular form of this technique is called *front-loaded* because it gives more weight to the most recent data (n) and reduces the significance of the older elements. For the front-loaded weighted moving average (see Figure 7.2)

$$w_1 \leq w_2 \leq \cdots \leq w_n$$

The weighting factors w_i may also be determined by regression analysis, but then they may not necessarily be front-loaded. A common modification to front-loading is called *step-weighting* in which each successive w_i differs from the previous weighting factor w_{i-1} by a fixed increment

$$c = w_i - w_{i-1}$$

The most common 5-day front-loaded, step-weighted average would have weighting factors increasing by $c = 1$ each day, $w_1 = 1$, $w_2 = 2$, $w_3 = 3$, $w_4 = 4$, and $w_5 = 5$. In general, for an n-day frontloaded step-weighted moving average:

$$w_{t-n+1} = 1$$
$$\vdots$$
$$w_{t-1} = n - 1$$
$$w_t = n$$

A TradeStation program for calculating an n-day, front-loaded, linearly weighted moving average is called *waverage*.

If simple linear step-weighting is not what you want, then a percentage relationship a between w_i elements can be used,

$$w_{i-1} = a \times w_i$$

If $a = 0.90$ and $w_5 = 5$, then $w_4 = 4.5$, $w_3 = 4.05$, $w_2 = 3.645$, and $w_1 = 3.2805$. Each older data item is given a weight of 90% of the more recent value. This is similar to *exponential smoothing* which will be discussed later in this chapter.

Weighting by Group

Prices may also be weighted in groups. If every two consecutive data elements have the same weighting factor, and p_t is the most recent price, n is the calculation period (preferably an even number), and there are $n/2$ number of weights, then

$$W_t = \frac{w_1 p_{t-n+1} + w_1 p_{t-n+2} + w_2 p_{t-n+3} + w_2 p_{t-n+4} + \cdots + w_{n/2} p_{t-1} + w_{n/2} p_t}{2 \times (w_1 + w_2 + \cdots + w_{n/2})}$$

For two or more data points using the same weighting, this formula can be regrouped as

$$W_t = \frac{w_1(p_{t-n+1} + p_{t-n+2}) + w_2(p_{t-n+3} + p_{t-n+4}) + \cdots + w_{n/2}(p_{t-1} + p_t)}{2 \times (w_1 + w_2 + \cdots + w_{n/2})}$$

Any number of consecutive data elements can be grouped for a step-weighted moving average.

    If the purpose of weighting is to reproduce a pattern that is intrinsic to price movement, then either the *geometric average*, $G = (p_1 \times p_2 \times p_3 \ldots p_n)^{1/n}$, discussed in Chapter 2, or *exponential smoothing*, explained later in this chapter, may be a better tool.

Triangular Weighting

While the simple moving average or linear regression treats each price equally, exponential smoothing and linear step-weighting put greater emphasis on the most recent data. There is an entire area of study in which the period of the dominant cycle is the basis for determining the best trend period. *Triangular weighting* or *triangular filtering*[1] attempts to uncover the trend by reducing the noise in both the front and back of the calculation window, where it is expected to have the greatest interference. Therefore, if a 20-day triangular weighting is used, the 10th day will have the greatest weight, while days 1 and 20 will have the smallest.

    To implement triangular weighting, begin with the standard formula for a weighted average, calculated for n days as of the current day t,

$$W_t = \frac{\sum_{i=1}^{n} w_i P_{t-n+1}}{\sum_{i=1}^{n} w_i}$$

[1]J. J. Payne, "A Better Way to Smooth Data," *Technical Analysis of Stocks & Commodities* (October 1989).

where n is also called the size, or width, of the window. For triangular weighting, the weighting factors w_i will increase linearly from 1 to the middle of the window, at $n/2$, then decrease to the end at n. This has a slightly different form when the period is odd or even,

$$
\begin{aligned}
w_i = i, & \quad \text{for } i = 1 \text{ to } int((n + 2)/2) \\
n - i + 1, & \quad \text{for } i = int((n + 2)/2) + 1 \text{ to } n \text{ (even values of } n) \\
n - i, & \quad \text{for } i = int((n + 2)/2) + 1 \text{ to } n \text{ (odd values of } n)
\end{aligned}
$$

where int is a function that returns the integer portion by truncation. For odd values of n, the weighting factor has the value i, where i ranges from 1 to $n/2$ (rounded up with the help of the function int) and the value of $n - i$ from $n/2$ to n. Instead of a triangular filter, which climbs in equal steps to a peak at the middle value, a *Gaussian filter* can be used, which weights the data in a form similar to a bell curve. Here, the weighting factors are more complex, but the shape of the curve may be more appealing,

$$
w_i = 10^x \quad and \quad x = \frac{3}{2} \times \left(1 - \frac{2i}{n} \right)^2
$$

Triangular weighting is often used for cycle analysis. Two techniques that use this method successfully, Hilbert and Fischer transforms, can be found in Chapter 11.

Pivot-Point Weighting

Too often we limit ourselves by our perception of the past. When a weighted moving average is used, it is normal to assume that all the weighting factors should be positive; however, that is not a requirement. The *pivot-point moving average* uses reverse linear weights (e.g., 5, 4, 3, . . .) that begin with a positive value and continue to decline even when they become negative.[2] In the following formula, the *pivot point*, where the weight is zero, is reached about $\frac{2}{3}$ through the data interval. For a pivot-point moving average of 11 values, the eighth data point is given the weight of 0:

$$
\begin{aligned}
PPMA_t(11) = (&-3p_{t-10} - 2p_{t-9} - 1p_{t-8} + 0p_{t-7} + 1p_{t-6} + 2p_{t-5} \\
&+ 3p_{t-4} + 4p_{t-3} + 5p_{t-2} + 6p_{t-1} + 7p_t) / 22
\end{aligned}
$$

The intent of this pattern is to reduce the lag by front-loading the prices. The divisor is smaller than the usual linear weighted average (where the sum of 1 through 11 is 66) because it includes negative values. The general formula for an n-day pivot-point moving average is[3]

$$
PPMA_t(n) = \frac{2}{n(n+1)} \sum_{i=1}^{n} (3i - n - 1) P_i
$$

[2]Patrick E. Lafferty, "End-Point Moving Average," *Technical Analysis of Stocks & Commodities* (October 1995).

[3]Don Kraska, "Letters to S&C," *Technical Analysis of Stocks & Commodities* (February 1996), 12.

FIGURE 7.3 S&P continuous futures, April through December 2010, with examples of the simple moving average, standard deviation average, geometric average, and exponential smoothing, all with calculation periods of 40 days.

A computer program and indicator that calculates and displays the pivot-point moving average, both called *TSM Pivot Point Average*, are available on the Companion Website. The negative weighting factors actually reverse the impact of the price move for the oldest data points rather than just give them less importance. For a short interval this can cause the trendline to be out of phase with prices. This method seems best when used for longer-term cyclic markets, where the inflection point, at which the weighting factor becomes zero, is aligned with the cyclic turn or can be fixed at the point of the last trend change.

Standard Deviation Moving Average

A unique trend calculation technique uses the standard deviation of prices.[4] This method creates a comparatively smooth trendline, *StdAvg*, by modifying the moving average value with a percentage of the standard deviation of prices. The following instructions are from the program *TSM Stdev Mvg Avg*, available on the Companion Website. It uses 5% of a 30-period standard deviation, and a 15-period moving average; however, each of these values can be changed. The result is shown for S&P futures in Figure 7.3.

```
      SD = StdDev(close,30);
     SDV = (SD – SD[1]) / SD;
  StdAvg = Average(close,15) + .05*SDV;
```

---

[4]Robert T. H. Lee, *Power Tools for Traders* (Hong Kong: MegaCapital Limited, 1997).

GEOMETRIC MOVING AVERAGE

The *geometric mean* is a growth function that is very applicable to long-term price movement. It was introduced in Chapter 2. It is especially useful for calculating the components of an index. The geometric mean can also be applied to the most recent n points at time t to get a *geometric average* similar in function to a moving average

$$G_t = \left(p_{t-n+1} \times p_{t-n+2} \times \cdots \times p_{t-1} \times p_t \right)^{(1/n)} = \left(\prod_{i=1}^{n} p_{t-i+1} \right)^{(1/n)}$$

The daily calculation is more complicated but, as shown in Chapter 2, could be rewritten as

$$\ln G_t = \frac{\ln p_{t-n+1} + \ln p_{t-n+2} + \cdots + \ln p_{t-1} + \ln p_t}{n}$$

$$= \frac{1}{n} \left(\sum_{i=1}^{n} \ln p_{t-i+1} \right)$$

This is similar in form to the standard moving average based on the arithmetic mean and can be written in either spreadsheet or program code as

```
GA = average(log(price),n)
```

Note that some software will use the function *log* although the calculation is actually the *natural log*, ln. Other programs will allow the choice of *log(value)* or *ln(value)*. A weighted geometric moving average, for n days ending at the current day t, would have the form

$$\ln G_t = \frac{w_1 \ln p_{t-n+1} + w_2 \ln p_{t-n+2} + \cdots + w_{n-1} \ln p_{t-1} + w_n \ln p_t}{w_1 + w_2 + \cdots + w_n}$$

$$= \frac{\displaystyle\sum_{i=1}^{n} w_i \ln p_{t-i+1}}{\displaystyle\sum_{i=1}^{n} w_i}$$

The geometric moving average itself would give greater weight to lower values without the need for a discrete weighting function. This is most applicable to long data intervals where prices had a wide range of values. In applying the technique to recent index or stock prices, this distinction is not as apparent. For example, if the historical index values vary from 10 to 1000, the simple average of those two values is 505, and the geometric average is 100, but for the three sequential prices of 56.20, 58.30, and 57.15, the arithmetic mean is 57.2166, and the geometric is 57.1871. A 5-, 10-, or 20-day moving average of stock prices, compared to geometric averages of the same intervals, show negligible differences. The geometric moving average is best applied to long-term historic data with wide variance, using yearly or quarterly average prices.

ACCUMULATIVE AVERAGE

An *accumulative average* is simply the long-term average of all data, but it is not practical for trend following. One drawback is that the final value is dependent upon the start date. If the data have varied around the same price for the entire data series, then the result would be good. It would also be useful if you are looking for the average of a ratio over a long period. Experience shows that price levels have changed because of inflation or a structural shift in supply and/or demand, and that progressive values fit the situation best.

RESET ACCUMULATIVE AVERAGE

A *reset accumulative average* is a modification of the accumulative average and attempts to correct for the loss of sensitivity as the number of trading days becomes large. This alternative allows you to reset or restart the average whenever a new trend begins, a significant event occurs, or at some specified time interval, for example, at the time of quarterly earnings reports or at the end of the current crop year.

DROP-OFF EFFECT

Many *rolling trend calculations* are distinguished by the *drop-off effect*, a common way of expressing the abrupt change in the current value when a significant older value is dropped from the calculation. Simple moving averages, linear regressions, and weighted averages all use a fixed period, or window, and are subject to this. For an *n*-period moving average, the importance of the oldest value being dropped off is measured by the difference between the new price being added to the calculation, *t*, and the one being removed, *t–n*, divided by the number of periods,

$$Drop\text{-}off\ effect = (p_t - p_{t-n})/n$$

A front-weighted average, in which the oldest values have less importance, reduces this effect because older, high volatility data slowly become a smaller part of the result before being dropped off. Exponential smoothing, discussed next, is by nature a front-loaded trend that minimizes the drop-off effect as does the average-off method.

EXPONENTIAL SMOOTHING

Exponential smoothing may appear to be more complex than other techniques, but it is only another form of a weighted average. It can also be more accurately called *percentage smoothing*, and has the added advantage of only needing the current price, p_t, the last *exponentially smoothed* value E_{t-1}, and the smoothing constant *a* (a percentage), to

compute the new value. The technique of exponential smoothing was developed during World War II for tracking aircraft and missiles and projecting their positions: the immediate past is used to predict the immediate future.

The terms of the geometric progression

$$a^0, a^1, a^2, a^3, \dots, a^n$$

become the weighting factors of a weighted moving average

$$W_t = \frac{w_1 p_{t-n+1} + w_2 p_{t-n+2} + \dots + w_{n-1} p_{t-1} + w_n p_t}{w_1 + w_2 + \dots + w_n}$$

in reverse order, where w_n is assigned the value a, $w_{n-1} = a^2$, $w_{n-2} = a^3$, and so forth. If $a = \frac{1}{2}$, which is also called *50% smoothing*, the resulting sequence of weighting factors is

$$1, \tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \dots, (\tfrac{1}{2})^n$$

where $a^0 = 1$ and $a^1 = a$. This shows the rapidly decreasing importance of each older price. Substituting the geometric progression into the equation for the weighted moving average gives the exponential solution

$$E_t = \frac{a^n P_{t-n+1} + a^{n-1} P_{t-n+2} + \dots + a^2 P_{t-1} + a p_t}{a^n + a^{n-1} + \dots + a^2 + a}$$

The Common Form of Exponential Smoothing

The common use of exponential smoothing takes a much simpler form,

$$E_t = E_{t-1} + a\,(p_t - E_{t-1})$$

where E_t and E_{t-1} = today's and yesterday's exponential smoothing values
  p_t = today's price
  a = the smoothing constant, $0 \leq a \leq 1$.

It may be easier to visualize the effect of the smoothing process with $a = 0.10$ by thinking of it as moving the exponential trendline closer to the current price by 10% of the distance between the price and the previous trendline value. In Figure 7.4, using a smoothing constant of 0.10, the distance from the new price p_t and the previous trendline value E_{t-1} is $p_t - E_{t-1}$. The new trendline value is 10% closer to the price because the distance between p_t and E_{t-1} is reduced by $0.10 - (p_t - E_{t-1})$. Therefore,

> *New exponential value = Previous exponential value*
> *+ a % of (Today's price – Previous exponential value)*

The smoothing process is started by letting $E_1 = p_1$ and calculating the next value:

$$E_2 = E_1 + a \times (p_2 - E_1)$$
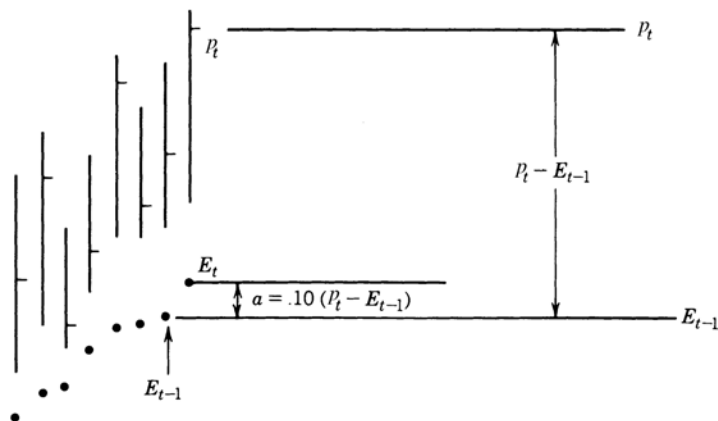
FIGURE 7.4 Exponential smoothing. The new exponential trendline value, E_t, is moved closer to the new price, p_t, by 10% of the distance between the new price and the previous exponential trendline value E_{t-1}.

Even though the calculations are initialized with the closing price, a longer-term smoothing, where the constant is small, will take more data for the smoothed line, E_t, to reach a stable value.

The Smoothing Constant Expressed in Days

The standard conversion from the number of days to a smoothing constant c was given by Hutson[5] as

$$c = \frac{2}{n+1}$$

where n is the equivalent number of days in the standard (linearly weighted) moving average. A, *2nd-* or *3rd-order exponential smoothing*, based on the weighting of the past 2 or 3 days' prices, is the exponential equivalent of step-weighting. Its general form is

$$c_p = 1 - \left(1 - \frac{2}{n+1}\right)^{\frac{1}{p}}$$

A comparison of the standard moving average days with 1st-, 2nd-, and 3rd-order exponential smoothing is shown in Table 7.3.

Using the Hutson conversion does not give you the real exponential smoothing equivalent of the moving average because it ignores the lag, which can be substantial. This can be seen in Tables 7.3 and 7.4 as well as in Tables 7.5 and 7.6. Most software development

---

[5]Jack K. Hutson, "Filter Price Data: Moving Averages vs. Exponential Moving Averages," *Technical Analysis of Stocks & Commodities* (May/June 1984).

TABLE 7.3 Comparison of Exponential Smoothing Values

Moving Average Days (*n*)	1st-Order *p* = 1	2nd-Order *p* = 2	3rd-Order *p* = 3
3	0.500	0.293	0.206
5	0.333	0.184	0.126
7	0.250	0.134	0.091
9	0.200	0.106	0.072
11	0.167	0.087	0.059
13	0.143	0.074	0.050
15	0.125	0.065	0.044
17	0.111	0.057	0.039
19	0.100	0.051	0.035
21	0.091	0.047	0.031

TABLE 7.4 Comparison of Exponential Smoothing Residual Impact

N	2/(*n* + 1)	RI (%)	10% RI	5% RI
5	0.333	13.17	0.369	0.451
10	0.182	13.44	0.206	0.259
15	0.125	13.49	0.142	0.181
20	0.095	13.51	0.109	0.139

platforms only allow the input of *days* for the calculation period, rather than the smoothing constant itself. This results in smoothing constants that have large jumps equivalent to the integer number of days. The program *TSM Exponential Smoothing*, available on the Companion Website, allows the specific smoothing constant to be entered.

Estimating Residual Impact

The primary difference between the standard moving average and exponential smoothing is that all prices remain as part of the exponentially smoothed value indefinitely. For practical purposes, the effect of the oldest data may be very small. A general method of approximating the smoothing constant c for a given level of residual impact is given by

$$c = 1 - RI^{1/n}$$

where n = the number of moving average days that is equivalent to the smoothing constant

    RI = the level of residual impact expressed as a percentage (e.g., 0.05, 0.10, 0.20).

A lower percent implies more residual impact.[6]

---

[6]Donald R. Lambert, "Exponentially Smoothed Moving Averages," *Technical Analysis of Stocks & Commodities* (September/October 1984).

TABLE 7.5 Evaluation of Exponential Smoothing—Significance of Prior Data

Weighting (%)											Past Number of Days												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1.00	100.0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.90	90.0	9.0	.9	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.80	80.0	16.0	3.2	.6	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.70	70.0	21.0	6.3	1.9	.6	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.60	60.0	24.0	9.6	3.8	1.5	.6	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.50	50.0	25.0	12.5	6.3	3.1	1.6	.8	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.40	40.0	24.0	14.4	8.6	5.2	3.1	1.9	1.1	.7	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.30	30.0	21.0	14.7	10.3	7.2	5.0	3.5	2.5	1.7	1.2	.8	—	—	—	—	—	—	—	—	—	—	—	—
.20	20.0	16.0	12.8	10.2	8.2	6.5	5.2	4.2	3.3	2.7	2.1	1.7	1.4	1.1	.9	—	—	—	—	—	—	—	—
.10	10.0	9.0	8.1	7.3	6.6	5.9	5.3	4.8	4.3	3.9	3.5	3.1	2.8	2.5	2.3	2.1	1.8	1.7	1.5	1.3	1.2	1.1	1.0

TABLE 7.6 Evaluation of Exponential Smoothing—Significance of Prior Data

Weighting (%)											Total Inclusion through nth Day												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1.00	100.0	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.90	90.0	99.0	99.9	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.80	80.0	96.0	99.2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.70	70.0	91.0	97.3	99.2	99.8	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.60	60.0	84.0	93.6	97.4	99.0	99.6	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.50	50.0	75.0	87.5	93.8	96.9	98.5	99.3	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.40	40.0	64.0	78.4	87.0	92.2	95.3	97.2	98.3	99.0	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.30	30.0	51.0	65.7	76.0	83.2	88.2	91.8	94.2	96.0	—	—	—	—	—	—	—	—	—	—	—	—	—	—
.20	20.0	36.0	48.8	59.0	67.2	73.8	79.0	83.2	86.6	89.3	91.4	93.1	94.5	95.6	96.5	—	—	—	—	—	—	—	—
.10	10.0	19.0	27.1	34.4	41.0	46.8	52.2	57.0	61.3	65.1	68.6	71.8	74.6	77.1	79.4	81.5	83.3	85.0	86.5	87.8	89.1	90.1	91.1

The approximation for the smoothing constant, given in the previous section as $2/(n + 1)$, can be shown to have a consistent residual impact of between 13 and 14%. The use of the preceding formula, as shown in Table 7.4, would allow the specific adjustment of residual impact; however, eliminating all the residual impact would transform exponential smoothing to a weighted average.

Relating Exponential and Standard Averages

It is much easier to visualize the amount of smoothing in a 10-day moving average than with a "10%" exponential smoothing, where $a = 0.10$. Although we try to relate the speed of both techniques, the simple moving average is equally weighted and the exponential is front weighted; therefore, they produce very different results. Because the exponential smoothing never completely discards the old data, a 10% smoothing is slower than a 10-day moving average, and a 5% smoothing is slower than a 20-day moving average.

To illustrate the difference in the trend patterns, a series of numbers from 1 through 15 and back to 1 will be used to compare a 5-day moving average with an "equivalent" 20% exponential smoothing. Figure 7.5 and Table 7.7 show the relationship between the moving average and exponential smoothing. During the period of constant increase and decrease of at least 5 consecutive days, the 5-day moving average stabilizes and follows the price series with constant lag of 1, equal to the price change (seen in column 4 of Table 7.7). During the same period of constant rise the exponential smoothing approaches the lag of 1 but only gets as close as 0.95. Had the steady rise been longer, it could have "theoretically" reached 1. The increased lag due to the residual values can be seen when the exponential trend changes direction four prices after the peak. The moving average changes direction three prices after the peak. In Figure 7.5 it is clear that the standard moving average has no memory of the previous price moves. The exact values can be seen in Table 7.7.
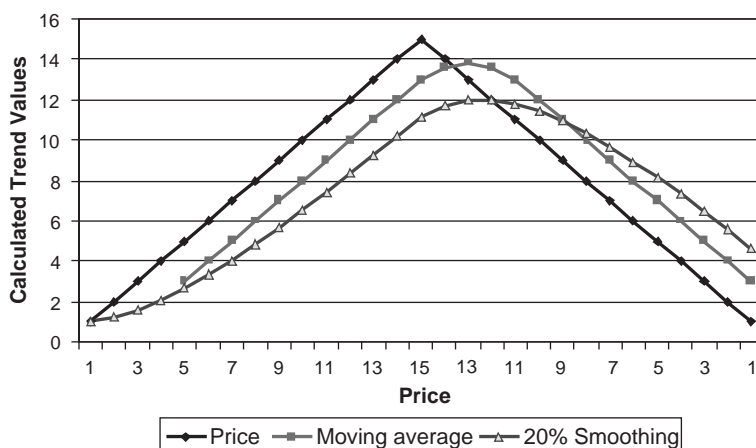
FIGURE 7.5   Comparison of trend lag. Exponential smoothing never quite reaches a constant lag and changes direction one period after the moving average trendline.

TABLE 7.7 Comparison of Lag between Standard and Exponentially Smoothed Averages

Price	Standard 5-Day	5-Point 20% Exp	Standard 20% Exp	Change 5-Day	Change 5–20% Exp	Change 20% Exp
1	—	1	1	—	—	—
2	—	1.2	1.2	—	—	—
3	—	1.56	1.56	—	Not enough	—
4	—	2.05	2.05	—	data	—
5	3	2.64	2.64	—	—	—
6	4	3.64	3.31	1	1	0.67
7	5	4.64	4.05	1	1	0.74
8	6	5.64	4.84	1	1	0.79
9	7	6.64	5.67	1	1	0.83
10	8	7.64	6.54	1	1	0.87
11	9	8.64	7.43	1	1	0.89
12	10	9.64	8.34	1	1	0.91
13	11	10.64	9.27	1	1	0.93
14	12	11.64	10.22	1	1	0.95
15	13	12.64	11.17	1	1	0.95
14	13.6	13.24	11.74	0.6	0.6	0.57
13	13.8	13.52	11.99	0.2	0.28	0.25
12	13.6	13.54	11.99	−0.2	0.02	0
11	13	13.36	11.79	−0.6	−0.18	−0.20
10	12	12.36	11.43	−1	−1	−0.36
9	11	11.36	10.95	−1	−1	−0.48
8	10	10.36	10.36	−1	−1	−0.59
7	9	9.36	9.69	−1	−1	−0.67
6	8	8.36	8.95	−1	−1	−0.74
5	7	7.36	8.16	−1	−1	−0.79
4	6	6.36	7.33	−1	−1	−0.83
3	5	5.36	6.46	−1	−1	−0.87
2	4	4.36	5.57	−1	−1	−0.89
1	3	3.36	4.65	−1	−1	−0.92

Figure 7.6 relates the fully calculated exponential smoothing (within 1%) to the standard moving average. Find the smoothing constant on the left scale, and the equivalent number of days in a standard moving average will be along the bottom. Observe in the following summary that, if you perform an optimization with equally spaced exponential smoothing constants, there is more sensitivity at the low end and little at the high end.

Equating Standard Moving Averages to Exponential Smoothing

Smoothing constant	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
Standard (*n*-day average)	20	10	6	4	3	2.25	1.75	1.40	1.15

A
Number of Days Used in Calculating
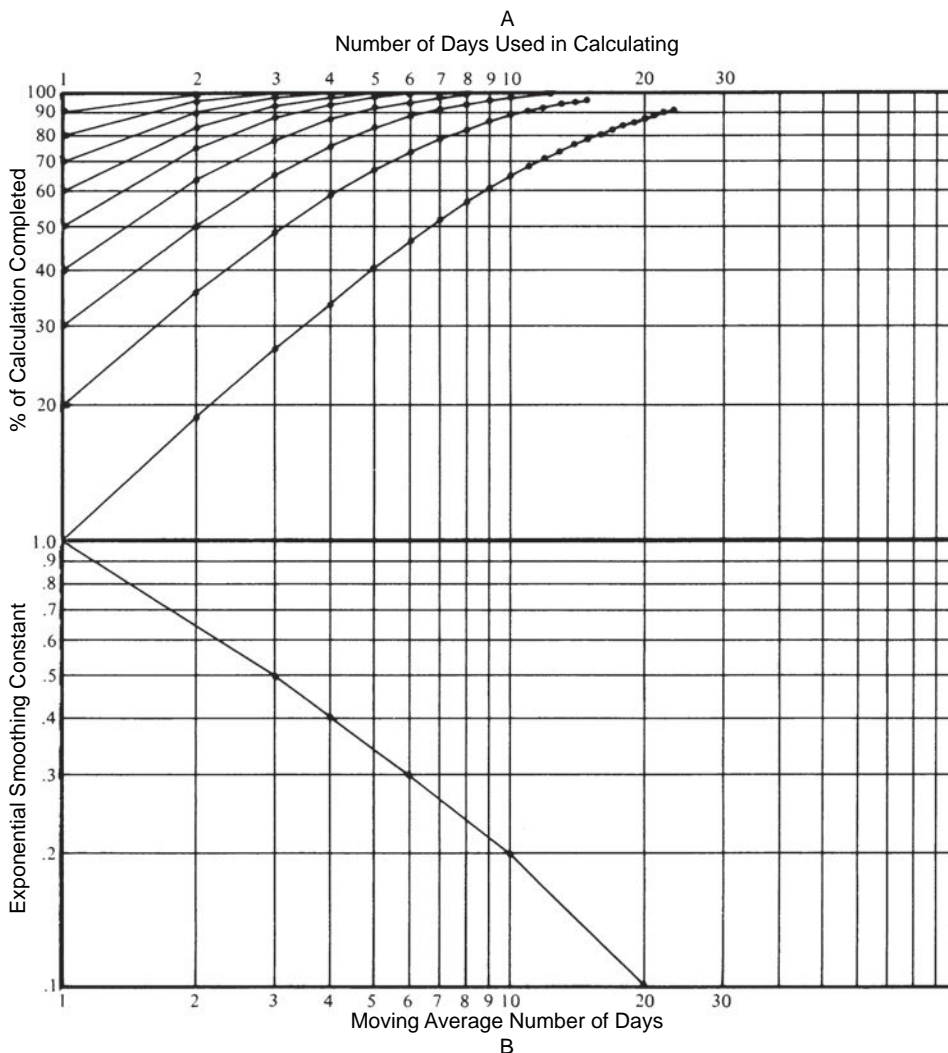
FIGURE 7.6   Graphic evaluation of exponential smoothing and moving average equivalents.

If equally distributed smoothing constants are used for testing, half of the tests will analyze moving averages of three days or less. If the process is reversed, and equally spaced days are used, then the smoothing constants are very different, shown in the next summary. It turns out that equally spaced smoothing constants is a better test distribution than equally spaced days. The distribution of calculation periods used for testing will be important when finding robust system parameters, and is discussed in Chapter 21.

Equating Exponential Smoothing to Standard Moving Averages

Standard (*n*-day average)	2	4	6	8	10	12	14	16	18	20
Smoothing constant	0.65	0.40	0.30	0.235	0.20	0.165	0.14	0.125	0.11	0.10

The distribution of smoothing constants is very close to logarithmic and is plotted on a log scale in Figure 7.6. This can be seen because the line representing the relationship between the smoothing constant and the moving average days, in the lower section, is nearly straight.

Correcting for the Lag

If the lag is considered the forecast error e_t in the exponential smoothing calculation, then

$$e_t = p_t - E_t$$

where E_t is the exponential trendline value (the smoothed value) corresponding to the most recent price p_t. The same smoothing technique can be applied to the lag, the pattern of increasing or decreasing errors, to get

$$ERR_t = ERR_{t-1} + a \times (e_t - ERR_{t-1})$$

The difference between the original smoothing value and this *2nd-order smoothing* of the errors ERR_t is added back into the approximation to get an adjusted trendline

$$EE_t = E_t + ERR_t$$

If the new error, the difference between the 1st-order exponential E_t and the 2nd-order approximation EE_t, is smaller than the original error e_t, then EE_t is an improvement in the forecast. This process can be continued to 3rd-order smoothing by applying the same method to the difference between the current price and the 2nd-order trend EE_t. This is similar to the method used by ARIMA discussed in Chapter 6.

Double Smoothing

In order to make a trendline smoother, the period of a moving average may be increased or the exponential smoothing constant decreased. This succeeds in reducing the short-term market noise at the cost of increasing the lag. An alternative to increasing the calculation period is *double smoothing*; that is, the trend values can themselves be smoothed. This will slow down the trendline, but gives weight to the previous values in a way that may be unexpected.

A double-smoothed 3-day moving average, *MA*, would take the most recent 3 moving average values, calculated from prices, and use them in another 3-period average to get a double-smoothed moving average, *DMA*:

$$MA_3 = (p_1 + p_2 + p_3)/3$$
$$MA_4 = (p_2 + p_3 + p_4)/3$$
$$MA_5 = (p_3 + p_4 + p_5)/3$$

then

$$DMA_5 = (MA_3 + MA_4 + MA_5)/3$$

By substituting the original prices, $p_1, \ldots, p_5$, into the equation for DMA_5 we find out that

$$DMA_5 = (p_1 + 2p_2 + 3p_3 + 2p_4 + p_5)/9$$

This shows that double smoothing puts weight on the center values and for a 3-day average the results look the same as *triangular weighting*. For longer calculation periods, the end values have decreasing weight but all other values have the same weight. For a 5-day double-smoothed average, the three end values would have declining weights. For exponential smoothing the result is also different. Because the most recent value in exponential smoothing receives the full weight of the smoothing constant, a, the double smoothing

$$DE_t = DE_{t-1} + a \times (E_t - DE_{t-1})$$

causes the nearby value t to be smoothed twice by a, or $a \times a$, and older values as well. Therefore, the net effect of using a constant of $a = 0.10$ for exponential double smoothing will result in a weighting that is much closer to using the square root of a, approximately 0.031.

Figure 7.7 and Table 7.8 give an example of this method using the standard conversion of a smoothing constant from the number of days. Prices for Microsoft are shown along with a single 0.20 smoothing, a double smoothing, and a single error correction. The error correction positions the trendline in the middle of the price move and may be a good candidate for mean reversion trading.
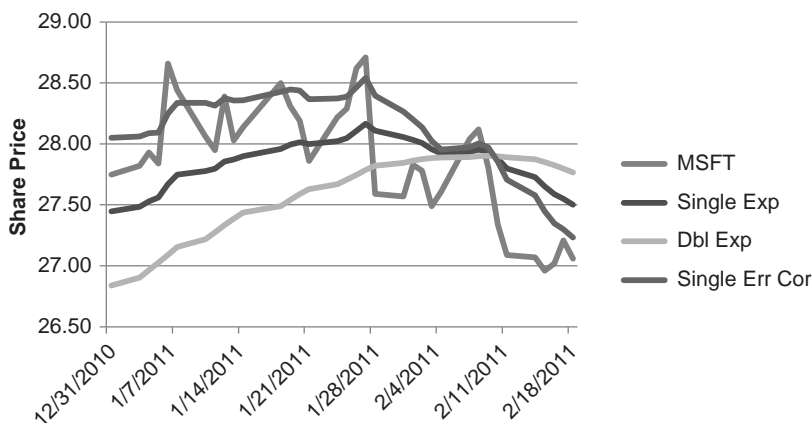
FIGURE 7.7 Comparison of exponential smoothing techniques applied to Microsoft shows a single 0.20 smoothing (*Exp*), a double-smoothed series (*Dbl Exp*), and a single smoothed series with the forecast error corrected (*Single Err Cor*).

TABLE 7.8 Comparison of Exponential Smoothing Techniques Applied to Microsoft

Date	MSFT	Exp 0.10	Dbl Exp 0.10	Exp Err	Sm Err 0.10	Corr Exp	Exp Chg	Dexp Chg
1/19/2011	28.31	27.994	27.54046	0.316	0.453338	28.44713		
1/20/2011	28.19	28.013	27.58775	0.177	0.425662	28.43908		Down
1/21/2011	27.86	27.998	27.62879	−0.138	0.369289	28.36736	Down	
1/24/2011	28.22	28.020	27.66793	0.200	0.352333	28.3726	Up	Up
1/25/2011	28.29	28.047	27.70586	0.243	0.341376	28.38862		
1/26/2011	28.62	28.105	27.74573	0.515	0.358787	28.4633		
1/27/2011	28.71	28.165	27.78766	0.545	0.377402	28.54247		
1/28/2011	27.59	28.108	27.81965	−0.518	0.287906	28.39546	Down	Down
1/31/2011	27.57	28.054	27.84307	−0.484	0.210735	28.26454		
2/1/2011	27.83	28.031	27.8619	−0.201	0.169519	28.20094		
2/2/2011	27.78	28.006	27.87634	−0.226	0.129939	28.13622		
2/3/2011	27.49	27.955	27.88417	−0.465	0.07048	28.02513		
2/4/2011	27.61	27.920	27.88777	−0.310	0.032414	27.9526		
2/7/2011	28.04	27.932	27.89221	0.108	0.039955	27.97212	Up	Up
2/8/2011	28.12	27.951	27.89809	0.169	0.052865	28.00382		
2/9/2011	27.81	27.937	27.90196	−0.127	0.034893	27.97175	Down	Down
2/10/2011	27.34	27.877	27.89948	−0.537	−0.02231	27.85486		
2/11/2011	27.09	27.798	27.88938	−0.708	−0.09093	27.70753		
2/14/2011	27.07	27.726	27.873	−0.656	−0.1474	27.57821		
2/15/2011	26.96	27.649	27.85061	−0.689	−0.20156	27.44749		

Double Smoothing of Price Changes

To reduce the compounded lag produced by double smoothing yet take advantage of the smoother trendline, William Blau has substituted the price changes, $p_t - p_{t-n}$, for the price itself,[7] a process that makes the data more sensitive to change. The first smoothing is then performed on this accelerated price series and acts to restore the speed of the series back to normal. When $n = 1$, the price changes are called the *first differences*, or *speed*. In effect, the first smoothed series does not have a lag so that the second smoothing results in one lag, the same as a normal moving average. Blau found this to be a successful proxy for long-term trends, where the first smoothing may be as long as 250 days, and the second a much shorter 5 days. The only disadvantage of this is that the scale is no longer the same as the price so that it cannot be plotted along with prices.

   To calculate the double smoothing of price change (in the following chapters the price change is also called *momentum*), follow the same method as exponential smoothing. First calculate the smoothed momentum, SM_t, substituting the price change, $p_t - p_{t-n}$, for the price, p_t, normally used. Next, perform another exponential smoothing using the smoothed momentum instead of price. The result is the double-smoothed momentum, DSM_t.

$$SM_t = SM_{t-1} + a((p_t - p_{t-n}) - SM_{t-1})$$

$$DSM_t = DSM_{t-1} + a(SM_t - DSM_{t-1})$$

---

[7] William Blau, *Momentum, Direction, and Divergence* (New York: John Wiley & Sons, 1995).
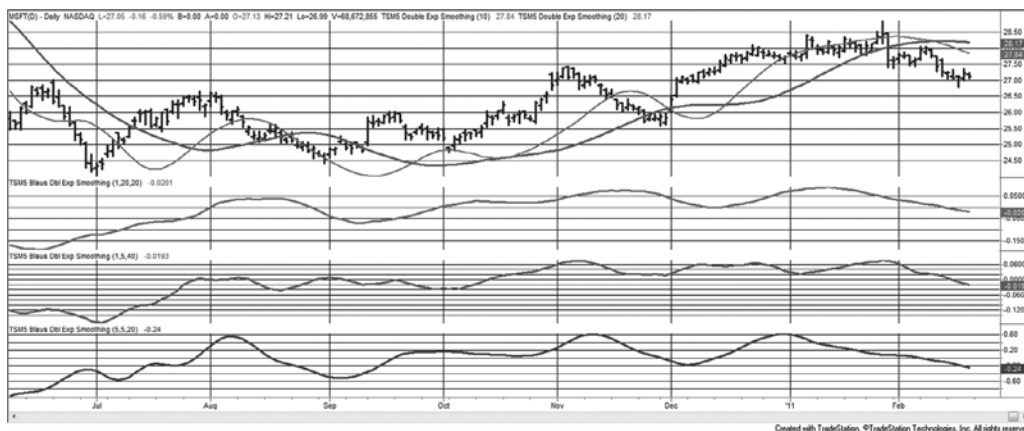
FIGURE 7.8 Double smoothing applied to Microsoft, June 2010 through February 2011. (Top panel) the price of Microsoft along with an exponential double smoothing of 10 and 20 days. (Second panel) Blau's double smoothing of 20/20 days. (Third panel) Blau's smoothing of 5/40 days. (Bottom panel) The 5-day difference before applying double smoothing of 5/20 days.

As an example of this approach, start with the 5-day momentum, $M(5)_t = p_t − p_{t–5}$, and smooth the momentum values using the 5-day period equivalent, 0.166. Smooth the resulting values again using the 20-day smoothing constant equivalent, 0.0555. The result is shown in the bottom panel of Figure 7.8 applied to Microsoft. The trendline is no longer in the same units as price; therefore, it cannot be plotted in the same panel as Microsoft prices. However, the double-smoothed trendline is very smooth, and the points at which the trends turn show less lag than any of the other methods. Trading signals should be generated from the trendline only, that is, buying when the trendline turns up and selling when it turns down. These methods are discussed in detail in Chapter 9 under the headings "True Strength Index" and "Double-Smoothed Stochastics."

Comparison of Exponential Smoothing Methods

The single and double exponential smoothing methods, including error corrections, can be entered into a spreadsheet and compared. A piece of that spreadsheet is shown in Table 7.8, and the entire spreadsheet is available on the Companion Website as *TSM Comparison of exponentials MSFT*. The methods are applied to Microsoft from August 1998 through February 2011. The first calculation in column 3 is the single exponential smoothing (*Exp*), followed by the double smoothing in column 4 (*DblExp*). The error for the single smoothing, the current price minus the corresponding exponential value, is in column 5 (*Exp Err*), followed by an exponential smoothing of those error values (*Sm Err*). The seventh column shows the single exponential with the smoothed errors

added back (*Corr Exp*). In order to see the difference in the single smoothing and the error-corrected method, the last two columns indicate when the trends turn up or down based on the trendline changing direction. The error-corrected method is most sensitive to changing prices.

Regularization

An interesting form of double smoothing is given by Mills and called *exponential regularization.*[8]

$$REMA_t = \frac{REMA_{t-1} \times (1 + 2 \times w) + \alpha \times (p_t - REMA_{t-1}) - w \times REMA_{t-2}}{1 + w}$$

where p = price
α = smoothing constant
w = weighting factor

Nominally, both α and w are set to 9. A comparison of a 9-day standard exponential and the 9-day exponential regularization is shown in Figure 7.9, applied to weekly S&P data, 2007 through 2010. The regularized trendline is much smoother. The program *TSM Exponential Regularization* can be found on the Companion Website.
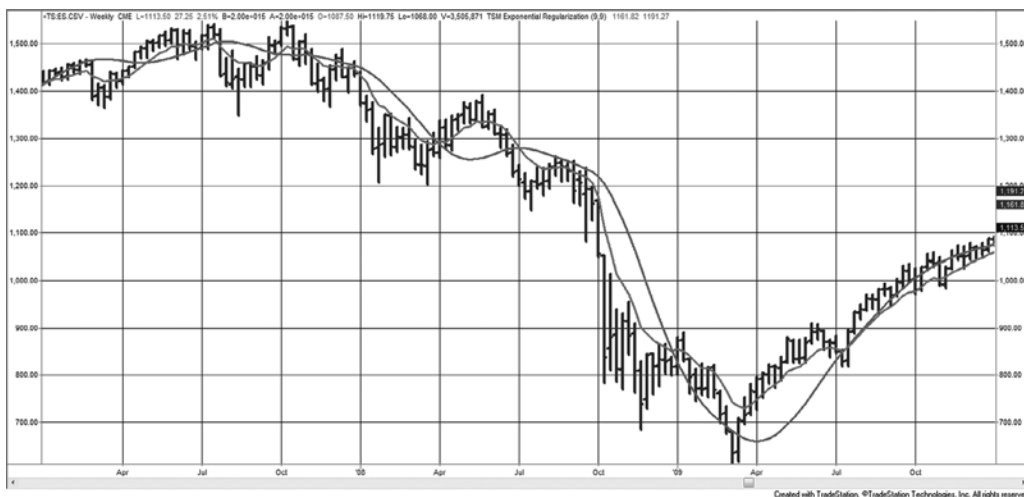
FIGURE 7.9 Comparison of a 9-day exponential smoothing with a 9-day exponential regularization (the smoother line), applied to the weekly S&P, 2007–2010.

[8]Mark Mills, "Regularization" (in "Traders' Tips"), *Technical Analysis of Stocks & Commodities* (July 2003).
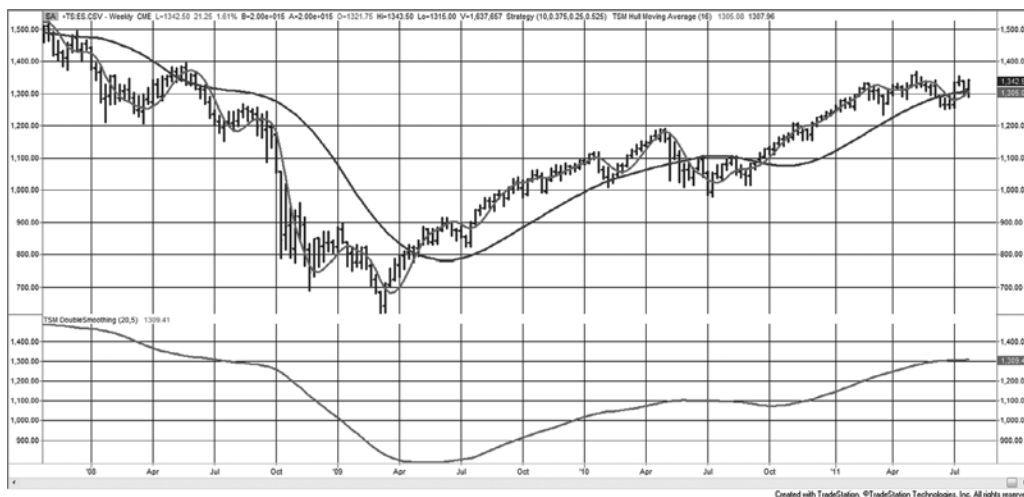
FIGURE 7.10   The Hull Moving Average (slower trend in the top panel), compared with a double-smoothed 16-week simple moving average (top panel) and a double-smoothed exponential (20- and 5-week) in the bottom panel.

Hull Moving Average

A double-smoothed method that uses linearly weighted averages and modified calculation periods, all applied to weekly data, is the Hull Moving Average (HMA).[9] Beginning with the suggested period, p, of 16 weeks, three weighted averages use the original period, the square root of the period, and half the period, with the goal of shorting the period and reducing the lag. The calculation can be written as

$$\text{WAVG1} = \text{WAVG (close, } p)$$

$$\text{WAVG2} = \text{WAVG (close, int}(p/2))$$

$$\text{HMA} = \text{WAVG}(2 \times \text{WAVG2} - \text{WAVG1, int(sqrt}(p)))$$

where $close$ = weekly closing prices
                 p = calculation period
         WAVG = weighted average function (data series, period)
            int = integer portion function
          sqrt = square root function

The 16-week HMA can be compared with a traditional double smoothing, using a 20-week exponential smoothing followed by a 5-week smoothing, as shown in Figure 7.10. Applied to the S&P *e*mini from December 2007 through July 2011, both of these are very

---

[9] "The Hull Moving Average," *The Technical Analyst*, (July–September 2010).

similar (the HMA is the slower trend in the top panel and the double smoothed average is at the bottom). A double-smoothed 16-week simple moving average is much more sensitive to price changes, as seen in the top panel. A program, *TSM Hull Moving Average*, can be found on the Companion Website.

PLOTTING LAGS AND LEADS

Moving averages, as well as any trend calculation, can be plotted with *leads* and *lags*, that is, forward or backward on a chart, each having a major impact on the interpretation. The conventional plot places the moving average value MA_t directly above or below the last price p_t used in the calculation. Any averaging or smoothing technique is said to be *lagging* actual price movement. When prices have been trending higher over the calculation period, the value MA_t will be lower than the most recent price p_t; therefore, it will be plotted below the actual prices. When prices are declining, the trendline will be plotted above the prices.

Using a simple moving average as an example, the trend values can be plotted so they *lead* or *lag* the most recent price used in the calculation. If it is to *lead* by 3 days, the value MA_t is plotted three periods ahead on the chart, on day $t + 3$; if it is to *lag* by 2 days, it is plotted at $t – 2$. In the case of leading moving averages, the analysis attempts to compensate for the time delay by treating the average as an n-day-ahead forecast rather than a concurrent statement of direction. A penetration of the forecasted line by the price may be used to signal a change of direction. The lag technique may serve the purpose of *phasing* the moving average, putting it in tune with a specific cycle. A 10-day moving average, when lagged by 5 days, will be plotted in the middle of the actual price data. This technique is covered later. Figure 7.11 shows Microsoft with a simple 1-day moving average plotted 2 and 4 days ahead as well as 2 and 4 days lagged.
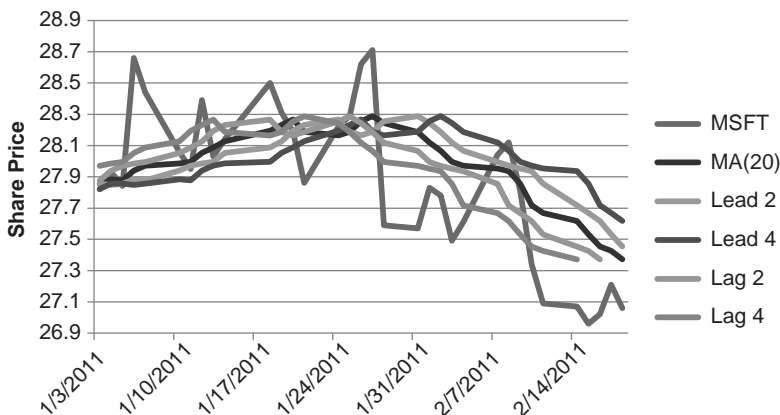
FIGURE 7.11   Plotting a moving average lag and lead for a short period of Microsoft.

Trend Systems

T he previous chapters developed the tools for calculating trends—a traditional moving average, various weighted averages, exponential smoothing, and regression. To profit from identifying the trend requires the use of trading rules and the selection of specific parameters that define the trend speed and an acceptable level of risk, among other factors. This chapter first discusses those rules that are necessary to all trading strategies and then gives examples of actual systems. The selection of trend speed is handled only briefly here but is continued with a detailed analysis of these and other systems throughout the book, and especially in Chapter 21. It is most important to find trends that are robust, that is, that work across many markets and under varied economic conditions. At the same time they must satisfy an investor's risk tolerance. It is a difficult balance.

Trend systems are the preferred choice of Commodity Trading Advisors (CTAs). Some advisors are reported to be using the same systems devised in 1980. Barclay Hedge (BarclayHedge.com) reports that hedge funds had a total of $1,762.9 billion under management as of the first quarter of 2012, and managed futures totaling $328.9 billion. Investments from both institutions and individuals have been flowing into the industry at an increasing rate, up almost 500% since 2002.

WHY TREND SYSTEMS WORK

Trend analysis is the basis for many successful trading programs, some with audited performance published for more than 30 years. Being able to identify the trend is also important if you are a discretionary trader looking to increase your chances for success by trading on "the right side of the market." Trend systems work because

- *Long-term trends capture large price moves caused by fundamental factors*. Economic trends are most often based on government interest rate policy, which is both slow to develop and persistent. In turn, interest rates directly affect foreign exchange, the trade balance, mortgage rates, carrying charges, and the stock market.

- *Prices are not normally distributed but have a fat tail*. The fat tail means that there is an unusually large number of directional price moves that are longer than would be expected if prices were randomly distributed. The fat tail will generate exceptional trend profits, which are essential to a trend system's long-term success.
- *Money moves the markets*. Most trends are supported by the flow of investor funds. While this causes short-term noise, it also delivers the long-term trends. As trends become clearer to the general public, additional money flows in to continue the trend.

Trend trading works when the market is trending. It doesn't work in markets that are not trending. There is no magic solution that will generate profits for trending strategies when prices are moving sideways, and there is no one trending technique that is always best. You'll find that most trending methods have about the same returns over time but with different risk profiles. It is the risk profile and the trading frequency that distinguish one method from another, and those features will be discussed throughout this chapter.

How Often Do Markets Trend?

Is there a way to measure how often markets trend? One analyst defines a trend as 10 consecutive closes in the same direction, but that seems arbitrary and a small window. What if there were nine days up and one small down day?

A trend is a relative concept. It is relative to the trader's time horizon, and it is relative to the amount of noise and price swings that are acceptable within the trending period. Ultimately, a trend exists if you can profit from the price moves using a trending strategy.

The Fat Tail

The *fat tail* is a statistical phenomenon caused by a combination of market fundamentals and supported by human behavior. The net effect is that prices move in one direction much longer than can be explained by a random distribution. As a simple example, consider coin flipping as a classic way to produce a random distribution. In 100 coin tosses,

- 50 will be a head or a tail followed by the opposite head or tail.
- 25 will be two heads or two tails in a row.
- 12½ (if we could have halves tosses) will be three heads or three tails in a row.
- About 6 will be four heads or four tails in a row.
- About 3 will be five in a row.
- 1 or 2 will be six in a row of either heads or tails.

If price moves are substituted for coin flips, then heads would be a move up and tails a move down. If the pattern of up and down price moves follows a random distribution (and the up and down moves were of the same amount), then it would not be

Random versus Actual Price Distribution
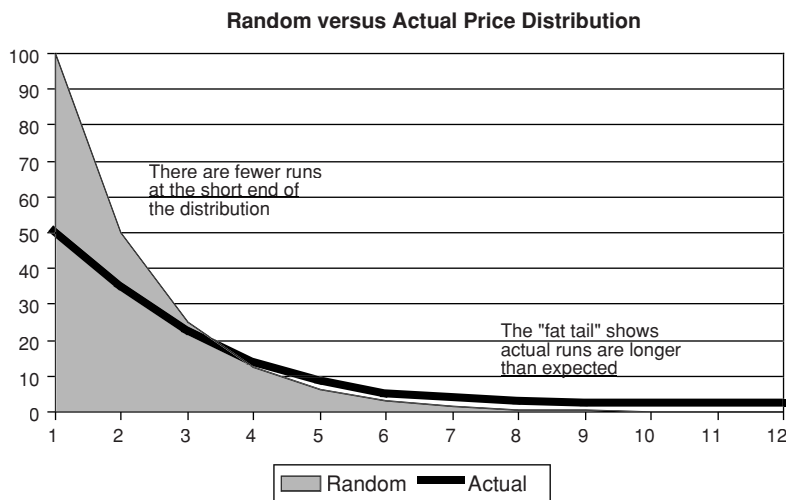
FIGURE 8.1   Distribution of runs. The shaded area shows the normal distribution of random runs. The solid dark line represents the distribution when there is a fat tail. In the fat tail distribution, there are fewer short runs and an unusually large number of longer runs or a single exceptionally long run.

possible to profit from a trend system. But prices are not normally distributed. Price runs have a fat tail, which means that, instead of one run of 6 out of 100 days of trading, we may see a run of 12, or 3 runs of 6. That distribution is enough to make trend trading profitable.

Another factor is that these long runs translate into very large trading profits. It is not necessary to have every day go in the same direction in order to profit, only that the downward reversals during an uptrend not be large enough to change the direction of the trend and end the trade. The more tolerance for the size of the interim reversals, the more likely the fat tail can be captured.

If there are more runs of longer duration for every 100 daily price moves, what is the shape of the rest of the distribution? Figure 8.1 gives a theoretical representation of an actual price distribution compared to a random distribution. The extra movement that goes into creating the fat tail comes from the frequency of short runs. There are fewer runs of 1 and 2 and more runs greater than 6. The total remains the same. Readers interested in this subject should read the section "Gambling Techniques—The Theory of Runs" in Chapter 22.

Distribution of Profits and Losses

As a trader, you would want to know "How often is there a profit from a fat tail?" To find the answer, we'll apply the most basic trending system, a simple moving average that buys when the trendline turns up and sells short when it turns down. This will be
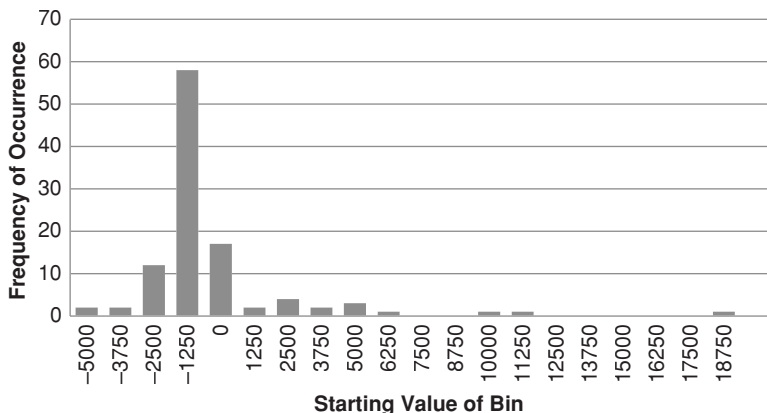
FIGURE 8.2 Frequency distribution of returns for S&P futures using a 40-day simple moving average strategy. Results show a fat tail to the right.

discussed in more detail in the next sections. For now we need to know that results depend on both the market and the calculation period. Applying a simple 40-day moving average strategy to five diverse futures markets, 30-year bonds, the S&P, the euro currency, crude oil, and gold, the results of individual trades can be collected and displayed as a histogram (frequency distribution). The results of the S&P are shown in Figure 8.2.

In the frequency distribution, the bottom axis shows the starting value of the bins that hold the size of the profitable or losing trades, and the left scale shows the number of trades that fall into that bin. If the distribution was normal, then the shape would be a bell curve. This distribution is clearly extended far to the right, with one very large profit showing in the $18750 bin. That one profit offsets more than 30 losses in the largest bar marked −$1250. But the S&P is not the only market with this distribution; actually, the S&P has relatively smaller fat tails than other markets. Table 8.1 shows the distribution sample. The tails to the right are very long and those to the left very short. It is important to remember that a pure trend strategy needs this distribution to be profitable.

TABLE 8.1 Frequency Distribution for a Sample of Five Diverse Markets, Showing the Fat Tail to the Right and a Short Tail to the Left

Bin	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Bonds	0	2	8	17	53	27	6	3	2	3	1	0	0	1	1	0	0	0	1	0
S&P	2	2	12	58	17	2	4	2	3	1	0	0	1	1	0	0	0	0	0	1
Euro	1	0	4	7	26	75	20	2	2	3	1	0	2	1	1	0	0	2	0	0
Crude	0	3	22	92	2	0	4	0	1	0	0	0	0	0	0	0	0	1	0	0
Gold	0	1	4	8	20	70	16	4	0	2	0	3	0	0	0	0	0	1	0	0

Time Intervals, Market Maturity, and Trends

Trends are most easily seen using long-term charts, weekly rather than daily data, or daily rather than hourly data. The farther you step back from a chart, the clearer the trend. If you display a daily chart, there will be some obvious trending periods and some equally clear sideways moves. Change that to a weekly chart, and the trends will seem much clearer. Change that to an hourly chart, and you'll see mostly noise. Lower frequency data translate into better performance when using longer-term trends. While there are always fast trends that show profits in backtesting, they tend to be less stable and inconsistent in their returns. Trends using longer calculation periods are more likely to track economic policy, such as the progressive lowering of interest rates by the central bank or a plan to allow the currency to weaken in order to stimulate exports and reduce debt.

Diverse markets may have very different trending qualities. Interest rate futures, money markets, and utility stocks are among the many investment vehicles closely tied to government rate policy and reflect the same long-term trend; this trend can persist for years. Foreign exchange is more complex and is constrained by monetary policy. Governments are more accepting of changes in the exchange rates if they occur slowly, but they will work hard to keep them within a target range. Most foreign exchange markets show clear but shorter trends compared to interest rates.

The stock market presents another level of difficulty. Individual stocks are driven by many factors, including earnings, competition, government regulation, management competence, and consumer confidence. Because the volume of trading in individual shares may vary considerably from day to day, these factors do not often net out as a clear trend. Stock prices may run up sharply on anticipation of better earnings or approval of a new drug, and reverse just as quickly within a few days. Liquidity, or volume, is an important element in the existence of a steady trend.

Individual stocks are also affected by concurrent trading in the index markets. When the S&P futures or the ETF SPY is bought and sold, all stocks within that index are bought and sold. If one company in the S&P 500 has just announced the loss of a major contract, but the overall market is strong, the share price of the suffering company may be dragged higher by arbitrage due to massive buying of the S&P index. This behavior makes for erratic price patterns in individual stocks.

Emerging markets are the exception. The introduction of a new market, such as the fictitious East European Stock Index, would be lightly traded but may be very trending. Initial activity would be dominated by commercials, such as banks, all of which would have a similar opinion on the economy of Eastern Europe. A small number of traders with the same opinion will cause very clear trends in the price. As the general public starts to participate, it adds liquidity while it also introduces noise, which in turn makes the trends less clear. Finding the trend then requires a longer time interval. Readers interested in this process should review the discussion of noise in Chapter 1.

When using a trending strategy, select both the markets and the time frame that work with you. Longer calculation periods, lower frequency data, and markets that are more closely linked to their underlying fundamentals will all perform better.

BASIC BUY AND SELL SIGNALS

Trends are based on the average always lagging price movement. It is the advantage and the disadvantage of the method. As the calculation period gets larger, the average lags further. Figure 8.3 shows Amazon prices from April 2010 through February 2011 with a 40-day moving average. Clearly, the moving average smoothes prices. The basic idea behind using the moving average as a trend signal is to be long when prices are above the average and short when below. The rules are stated as

- *Buy* when prices cross above the trendline.
- *Sell short* when prices cross below the trendline.

Even with these simple rules, there are important choices to be made. Do you buy at the moment rising prices cross the trendline during the trading session, or do you wait for the price to close above the trendline? As seen in Figure 8.3, prices may cross back and forth through the trendline before settling on a final direction. If you subscribe to the belief that the closing price of the day is the most reliable price, then the number of trading signals can be reduced by using the rules:

- *Buy* when prices close above the trendline.
- *Sell short* when prices close below the trend line.

Another school of thought prefers the average of high and low prices, or the average of the high, low, and closing prices. A buy or short sale signal occurs when the (high + low)/2 or (high + low + *close*)/3 crosses above or below the current trendline value. In both of these cases, the averages could not be calculated until the end of the trading session because none of the three component values would be known until then.

FIGURE 8.3   Amazon (AMZN) with a 40-day moving average.

Using the Trendline for Signals

The trendline represents the netting of all prices. Its purpose is to remove the price noise and show you the underlying price direction. Then it seems more reasonable to use the trendline to generate the trading signal. The change in the value of the trendline from the previous calculation to the current one is its direction—up, down, or unchanged. The direction of the trendline is a candidate for generating trading signals.

- *Buy* when the change in the trendline is up.
- *Sell short* when the change in the trendline is down.

The penalty for using the trendline as a trading signal is its lag. Figure 8.3 shows that, using a 40-day moving average, prices cross above the trendline during July a few days ahead of the point where the trendline turns up. The benefit using the trendline signal is that the trend turn is clear and more stable. During July and November, prices crossed back and forth through the trendline, but the trendline did not change direction; therefore, the trading signal remained the same.

Comparing Basic Trading Signals

The main differences between using a price penetration or the trendline to generate signals is that the trendline produces fewer trading signals, and those signals are delayed. If Amazon prices are used as an example, both methods of entry can be compared for a sample of calculation periods to see if one approach is consistently better than the other. Although this is a single example, it represents the general case.

As shown in Table 8.2, five calculation periods are used beginning with 5 days and doubling the period for each test. This maintains the percentage change in the calculation period and gives a better distribution sample (this is discussed further in Chapter 21). The two columns headed *Number of Trades* show that the trendline method has from

TABLE 8.2 Comparison of Entry Methods for 10 Years of Amazon (AMZN)

Signals using the trendline direction are shown on the left, and price penetration on the right.

| Trend Calculation Period | Signal Using Trendline | | | Signal Using Price Cross | | | |
	Total Profit/ Loss	Profit Factor	Number of Trades	Total Profit/ Loss	Profit Factor	Number of Trades	Increase in Trades
80	48.24	1.34	84	57.16	1.31	106	26%
40	94.42	1.46	120	32.21	1.12	164	37%
20	111.97	1.45	196	(7.31)	0.98	252	29%
10	(87.67)	0.81	292	(90.82)	0.82	370	27%
5	(90.31)	0.84	439	(49.15)	0.92	597	36%

26% to 37% fewer trades and, for the most part, better performance. The *Profit Factor* is the performance ratio used by TradeStation equal to the gross profits divided by the gross losses. While not as good as the *information ratio* or *Sharpe ratio*, the results allow you to compare performance.

It appears that using the trendline signals is always better, but that is not necessarily the case. To be certain, you would need to compare a wide range of diverse markets rather than just Amazon. Even if these results seem compelling, it is always necessary to confirm the numbers to be confident of the answers and to be sure that you understand the process.

While it seems convincing that the trendline is best for the slower trends, it is not as clear for faster trading. For the case using a 5-day moving average, both methods netted a loss, but the price penetration shows a smaller loss. It is very possible that, for faster trading, the lag in the trendline is too much of a burden to overcome and the price penetration is better. This will be discussed further in Chapter 16, Day Trading.

Anticipating the Trend Signal

Consistency is important. The system that is tested and the one that is traded should be the same. In this book, the closing price is used for most of the calculations; however, any combination of open, high, low, and close could be substituted. The normal process for generating a trading signal is to wait until prices close, then calculate the new moving average or trendline value, then see whether a crossing occurred or the direction of the trendline changed according to the basic buy and sell rules. But using the closing price for the calculation of the entry signal implies that you could enter on the close.

The process of waiting for the close price to perform the necessary calculations and generate a signal requires that orders be placed in the after-hours market or on the next open. While the trading system is indicating a new buy or sell signal as of the close of trading, you are entering the market significantly late. You are not following the system as it was tested. A practical solution to this dilemma is to record the prices shortly before the close, generate the trading signals, then enter the buy and sell orders for execution on the close. Occasionally, the order will be wrong because prices changed direction in the last few minutes of trading, but the cost of exiting the trade will usually be small compared to the improvement in overall execution.

The other alternative is to calculate, in advance, the closing price that will generate a signal using either the trendline method or the price crossing method. For an n-day moving average the calculation is simple—the new moving average value will be greater than the previous value if today's price is greater than the price dropped off n-days ago. Because all of the other values in the average remain the same except for the first and last, the answer only needs those two values. If a 40-day average is used and the oldest price p_{t-0} was 30.25, then any price greater than 30.25 today would cause the trendline to move up, and any price greater than 30.25 would also cause an upwards price penetration. Then an order can be placed in advance to *buy at 30.26 stop*.

How important is this? A lot depends on the trending nature of the market. In Chapter 1, the discussion of market noise showed that the short-term interest rates had the lowest amount of noise, and the equity index markets had the most noise. Using the Eurodollar interest rates and the S&P 500 futures as the extremes, the method that enters using the trendline was compared when entries were taken on the current close, the next open, and the next close to assess the sensitivity of the total profits to entry delays. Table 8.3 shows the results.

All Eurodollar results show a profit for trades entered on the close of the day that a signal occurred. Longer trend periods were generally more profitable, confirming the premise that trends are more dominant in the long term due to government policy. When entries were delayed to the next open or next close, the results were worse. This is best seen in the profit factors, which measure reward to risk rather than only total profits.

The S&P is also clear but not as clean as Eurodollars. Equity index markets are noisy; that is, prices moving up are not expected to continue up for any sustained number of days. A new buy signal will usually happen on a day when prices are moving higher, but waiting until the next open or the next close could be better because of the noise. This was true for all calculation periods except 5 days. As for the calculation period, there was no particular pattern, although absolute losses were smaller as the trend period got longer, a modest confirmation that longer trends are better. We know from other tests and other markets that this limited example is generally true, but market characteristics always create exceptions.

Profile of a Simple Moving Average System

Using the moving average trendline as the basis for system signals, we chose a 40-day calculation period because it tends to be the fastest one that also identifies the major price trends. The profile of results is typical of any moving average system. Figure 8.4 shows the NASDAQ 100 for one year, ending February 2011. Buy and sell signals are generated from the direction of the trendline; there were no transaction fees.

The trading signals in Figure 8.4 show that the major upwards move is captured, but not before there are a number of false signals due to the trendline changes during sideways price periods. However, given enough tries, the trend surfaces, and the system gains a large profit, similar to the fat tail discussed at the beginning of this chapter. The profile of this NASDAQ example, shown in Table 8.4, is typical of longer-term trend-following systems. Of the 150 trades over 10 years, only 52 of them were profitable, about 35%. However, the average winning trade was much larger than the average losing trade, with a ratio of 2.22, and winning trades were held much longer than losing trades, supporting the adage "cut your losses and let your profits run." Finally, there were more consecutive losses than consecutive profits, but that follows because there are many more losses. The performance picture is that trend following gets in and out quickly when it has a loss but holds the trade whenever trends and profits develop. This category of strategy is called *conservation of capital*, referring to the feature that cuts losses quickly.

TABLE 8.3 Comparison of Entries on the Close, Next Open, and Next Close

Results vary with the trending nature of the market. Analysis uses 10 years of S&P and Eurodollar interest rate futures, back-adjusted, ending in February 2011.

Calculation Period	Eurodollar Interest Rates						S&P 500					
	Today's Close		Next Open		Next Close		Today's Close		Next Open		Next Close	
	Total Profit or Loss	Profit Factor	Total Profit or Loss	Profit Factor	Total Profit or Loss	Profit Factor	Total Profit or Loss	Profit Factor	Total Profit or Loss	Profit Factor	Total Profit or Loss	Profit Factor
80	16150	2.87	16325	2.89	15630	2.75	(12325)	0.87	(10463)	0.89	16363	1.20
40	9603	1.61	9050	1.55	9383	1.61	(29138)	0.75	(26013)	0.78	5013	1.05
20	7745	1.37	6258	1.29	3850	1.16	13088	1.12	13900	1.13	22738	1.20
10	10773	1.40	8765	1.31	1165	1.04	(27925)	0.83	(22738)	0.86	(12550)	0.93
5	3368	1.09	(1715)	0.96	(870)	0.98	(21725)	0.90	(19588)	0.91	(563350)	−0.76
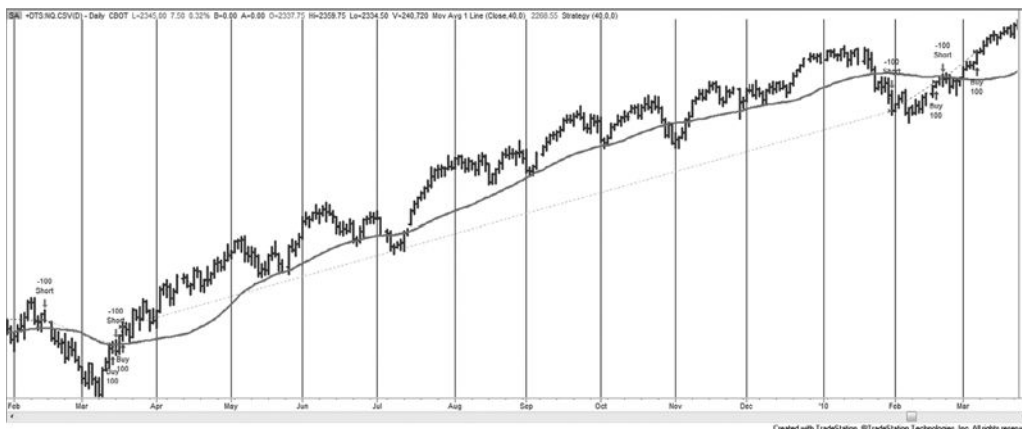
FIGURE 8.4 A trend system for NASDAQ 100. Applying a 40-day moving average and taking the trading signals from the direction of trendline results gives a typical trend system profile.

We can generalize the trend-following profile as:

- The percentage of profitable trades is low—often less than 30%.
- The average winning trade must be significantly larger than the average losing trade; actually, given only 30% profitable trades, the ratio must be greater than 100:30 to be profitable.
- The average winning trades are held much longer than losses.
- There is a high frequency of losing trades; therefore, there are also long sequences of losing trades.

There are many analysts that have a lifetime goal of improving these statistics, that is, capturing the long-term trend but improving the percentage of profitable trades. Some

TABLE 8.4 Performance Statistics for NASDAQ Futures, 10 Years Ending February 2011

Total profit	$11,880
Number of trades	150
Number of winning trades	52
Percentage of winning trades	34.7%
Average winning trade	1424.71
Average losing trade	−41.28
Win/loss ratio	2.22
Average bars in winners	31.65
Average bars in losers	6.64
Consecutive winner	6
Consecutive losses	10

small amount of success is possible but not without changing the risk characteristics of trend-following systems. For example, if you add profit-taking (discussed throughout the book) or stop-losses (Chapter 23), then you reduce or eliminate the chance of capturing the fat tail, which has been shown to be necessary for a long-term profit. Still, many traders do not like the idea of holding the trades for such a long time and giving back so much of the unrealized profits when the major trend changes direction. Different traders make different choices. A program to test the entry rules and execution options, including *long-only*, is *TSM Moving Average*, available on the Companion Website.

BANDS AND CHANNELS

A good way to improve the reliability of signals without altering the overall trend profile is by constructing a *band*, or *channel*, around the trendline. It can be used to effectively slow down trading without sacrificing the biggest profits. If we accept the premise that the point of trend change is also the time of greatest indecision, then a simple way to avoid frequent false signals is by using a band.

Bands Formed by Highs and Lows

The most natural band is one formed from the daily high and low prices. Instead of applying the n-day moving average to the closing prices, it is applied separately to the highs and lows. Long positions are entered when today's high crosses the average of the highs and short sales when today's low crosses the average of the lows. To get a broad view of whether this is an improvement to entry points, the two most extreme markets (the Eurodollar considered the trendiest and the S&P the noisiest) are tested for 10 years with the five calculation periods used in an earlier example. Results are shown in Table 8.5.

TABLE 8.5 Results of Using a Moving Average of the Highs and Lows, Compared to the Closes

	Eurodollar Interest Rates				S&P 500			
	Close Crossing		High-Low		Close Crossing		High-Low	
Calculation Period	Total Profit or Loss	Profit Factor	Total Profit or Loss	Profit Factor	Total Profit or Loss	Profit Factor	Total Profit or Loss	Profit Factor
80	16320	2.94	13842	2.18	30027	1.54	36443	1.74
40	16035	2.18	15000	2.02	(10337)	0.91	12402	1.14
20	10172	1.44	5167	1.20	5987	1.05	16512	1.15
10	2727	1.08	3667	1.11	(49000)	0.76	(33751)	0.82
5	7812	1.20	(337)	0.99	(106950)	0.64	(43760)	0.81

For a highly trending market, such as the Eurodollar interest rates, entering later, on a penetration of either the highs or lows, is not as good as entering on a price penetration of the close. If this had not been determined from the noise study in Chapter 1, these results would cause us to draw the same conclusion. Just the opposite is seen in the S&P results. Waiting longer to enter improves results noticeably, and, in the case of the 40-day trend, it turns a loss into a profit.

We can conclude that a band can be a profitable variation to a simple trend system, but not for all markets. The next question is "Are there other bands that work better?"

Keltner Channels

One of the original band calculations was by Keltner,[1] which goes as follows:

(Average daily price)	$AP_t = (H_t + L_t + C_t)/3$
(10-day moving average)	$MA_t = average(C_t, 10)$
(Upper band)	$UB_t = MA_t + AP_t$
(Lower band)	$LB_t = MA_t + AP_t$

These days we would tend to use true range, rather than the high-low range as a better representation of volatility.

Percentage Bands

Another simple construction is a percentage band, formed by adding and subtracting the same percentage of price from the trendline based on the closing prices. If c is the percentage to be used (where $c = 0.03$ means 3%), then

$$\text{(Upper band)} \quad B_U = (1 + c) \times MA_t$$
$$\text{(Lower band)} \quad B_L = (1 - c) \times MA_t$$

where MA_t = today's moving average value

Therefore, if the moving average value for Merck (MRK) is \$33, and the band is 3%, then the upper band is 33.99 and the lower band is 32.01. Because the moving average is much smoother than the price series, the band will be uniform around the moving average, narrowing and widening slightly as prices decline and rise.

The band can be more sensitive to change if the current price p_t is used to calculate the band instead of the moving average trendline. The bands are then

$$\text{(Upper band)} \quad B_U = (1 + c) \times p_{t-1} + MA_t$$
$$\text{(Lower band)} \quad B_L = (1 - c) \times p_{t-1} + MA_t$$

---

[1] Chester W. Keltner, *How to Make Money in Commodities* (Kansas City, MO: The Keltner Statistical Service, 1960).

The band is still oriented around the moving average trendline to prevent it from jumping up and down too often. Using the price to generate the band also requires that yesterday's price be used; otherwise the band would never be penetrated. Using a percentage band would work for stocks but not for futures. Most futures analysis uses back-adjusted data, and if the carry is mostly negative, then the oldest back-adjusted data could actually become negative. That doesn't affect a moving average trading signal, but it does make any percentage price calculation wrong. Back-adjusted historic prices are simply not correct.

If only a small buffer zone is needed, rather than one that adjusts over time, then a band based on absolute point value could be used. For example, stocks trading between $20 and $40 could have a $1 band. If your goal is to avoid a few very small variations around the time of trend change, then an absolute point value band may be a simple alternative solution.

Volatility Bands

To the degree that volatility increases as prices increase, the percentage band is a volatility band; however, that relationship is both long term and not as good as actually measuring the individual volatility for many stocks and futures markets. Regardless of the method, creating a band that is responsive to volatility may improve the reliability of trend signals. The independent smoothing of the high and low prices over any calculation period forms a natural volatility band. Although it may be practical to use the same smoothing technique or the same calculation period as the underlying trend (e.g., 10-day or 10% smoothing constant) for the high, low, and closing prices, it is not a requirement. If the same smoothing criterion is used, the band will be uniform with respect to the moving average of the closing price; if not, all three trendlines may weave around one another, which creates some practical problems.

There are many choices for measuring volatility and creating a band around the trendline. All of the methods of forming bands are subject to scaling. Scaling is accomplished by using a constant value as a multiplier or scaling factor; it increases or reduces the sensitivity of the band. If s is a scaling factor and c is a fixed percentage, then the following bands can be constructed:

$$B_t = MA_t \pm s \times c \times MA_t \qquad \text{(Percentage of trendline)}$$
$$B_t = MA_t \pm s \times c \times p_t \qquad \text{(Percentage of price)}$$
$$B_t = MA_t \pm s \times ATR_{t-1} \times MA_t \qquad \text{(Average true range)}$$
$$B_t = MA_t \pm s \times stdev_{t-1} \times MA_t \qquad \text{(Standard deviation)}$$

When $s = 1$, the scaling effect is nullified; for $s > 1$, the width of the band is increased; and for $s < 1$, the band width is reduced. In the choices above, MA was used to indicate a moving average, but any method of calculating the trend can be substituted, such as an exponential smoothing or a regression. Figure 8.5 shows the four types of bands applied to the S&P futures. All use a scaling factor of 2, which may be too close for some methods
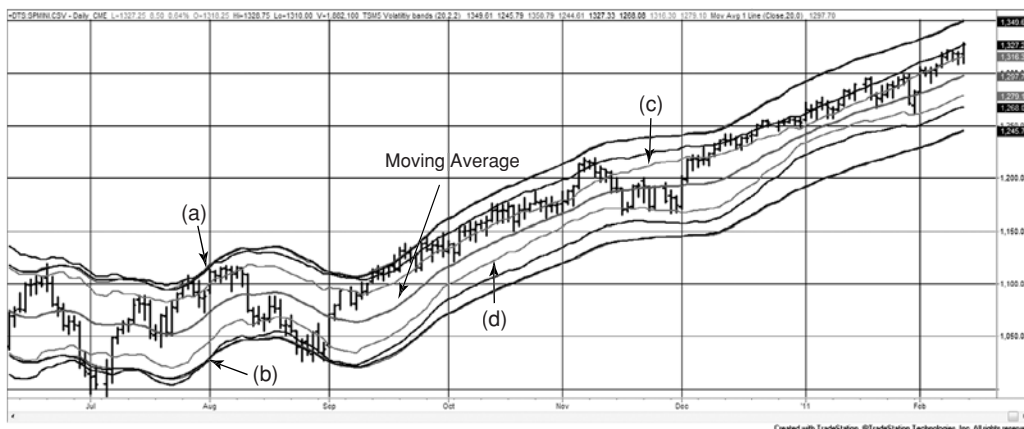
FIGURE 8.5 Four volatility bands around a 20-day moving average, based on, (a) 2% of trendline, (b) 2% of price, (c) 2 × average true range, and (d) annualized 20-day volatility.

and too far for others. The purpose of the chart is to show the relative shape of the bands and distance from the prices.

In Figure 8.5 the center line is a 20-day moving average. The first two methods of calculating bands, the percentage of trendline and percentage of price, are almost identical, very smooth, and are the farthest from the center. The next band closer to the moving average is the average true range. It moves slightly farther apart when prices are more volatile. The band closest to the trendline is the annualized volatility, which is most sensitive to price changes.

Because the same scaling factor produces bands that are different for each method, it is difficult to compare them without finding the scaling factors that come closest to average band width for each technique. The bigger decision is whether it is more sensible to use a very smooth band or one that reacts to changes in price volatility. That decision is up to each trader.

It may be convenient to have separate exit and entry bands, the entries less sensitive than the exits so that the strategy exits quickly but enters slowly. Or, if the entry occurs on a penetration of the band, but the exit is based on the trendline, then trades are not reversed from long to short. That improves slippage because only half the number of shares or contracts are traded on each order, and may avoid some false signals.

Bollinger Bands

Perhaps the simplest and most robust measurement of price volatility is the standard deviation of the prices themselves, calculated over recent price history. This was the last method listed in the previous section. John Bollinger has popularized the combination of a 20-day moving average with bands formed using 2 standard deviations of the price changes over the same 20 day period. They are now frequently called
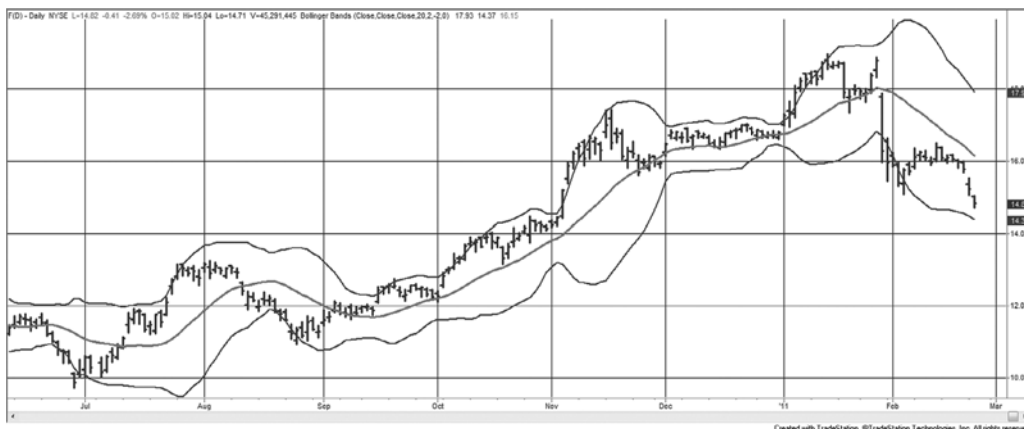
FIGURE 8.6   Bollinger bands applied to Ford.

Bollinger bands.[2] Because the standard deviation represents a confidence level, and prices are not normally distributed, the choice of two standard deviations equates to an 87% confidence band (if prices were normally distributed, two standard deviations would contain 95.4% of the data). In their normal use, Bollinger bands are combined with other techniques to identify extreme price levels. These are discussed later in this section.

Figure 8.6 shows Ford (F) plotted with a traditional Bollinger band. One of the characteristics of this band is that, once the price moves outside either the upper or lower band, it remains outside for a number of days in a row. This type of pattern was typical of what used to be called *high momentum*. Note that the width of the band varies considerably with the volatility of prices and that a period of high volatility causes a "bubble," which extends past the period where volatility declines. These features and more about volatility will be discussed in Chapter 20.

Figure 8.6 was created using the TradeStation indicator *Bollinger Bands*, which lets you vary both the calculation period for the trend and the number of standard deviations. But then, if it's not a 20-day average and 2 standard deviations, it's not a Bollinger band.

Bollinger bands can also be applied effectively to multiple time frames. An excellent example that uses a combination of weekly and daily data applied to the S&P 500 is seen in Figure 8.7. The price pattern follows the weekly Bollinger band higher, where the daily and weekly prices come together during the week of July 14.

Modified Bollinger Bands

One of the significant problems with Bollinger bands, as well as any volatility measure based on historic data, is that the bands will expand after increasing volatility but are

---

[2]John A. Bollinger, Bollinger Capital Management, Inc. P.O. Box 3358, Manhattan Beach, CA 90266, www.bollingerbands.com. Also *Bollinger on Bollinger Bands* (New York: McGraw-Hill, 2001).
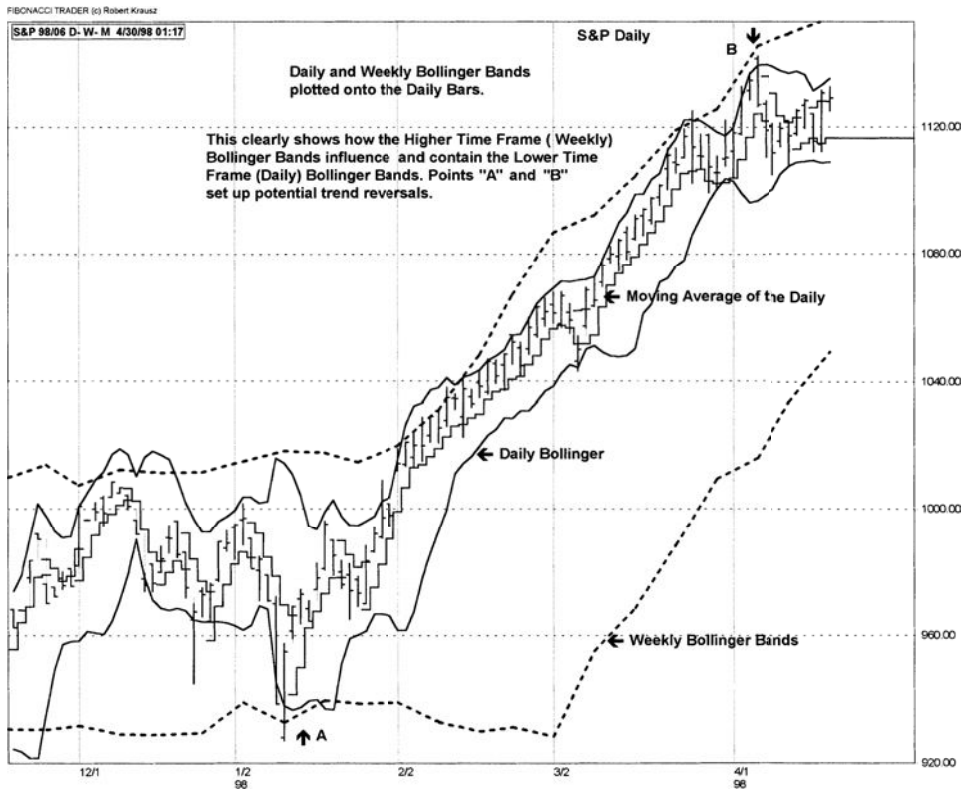
FIGURE 8.7 Combining daily and weekly Bollinger bands.

Source: Chart created using The Fibonacci Trader by Robert Krausz. Used with permission from Fibonacci Trader Corporation, St. Augustine, FL. www.fibonaccitrader.com.

slow to narrow as volatility declines. An excellent correction[3] for this requires the following calculations for the center line, D,

$$M_t = \alpha \times C_t + (1 - \alpha) \times M_{t-1}$$
$$U_t = \alpha \times M_t + (1 - \alpha) \times U_{t-1}$$
$$D_t = \frac{(2 - \alpha) \times M_t - U_t}{1 - \alpha}$$

where C is the closing price and α is the smoothing constant, set to 0.15 to approximate a 20-day moving average. In order to correct the bulge in the bands following a volatile period, the upper and lower bands (BU and BL) are calculated as

$$m_t = \alpha \times |C_t - D_t| + (1 - \alpha) \times m_{t-1}$$
$$u_t = \alpha \times m_t + (1 - \alpha) \times u_{t-1}$$

---

[3]Dennis McNicholl, "Better Bollinger Bands," *Futures* (October 1998).
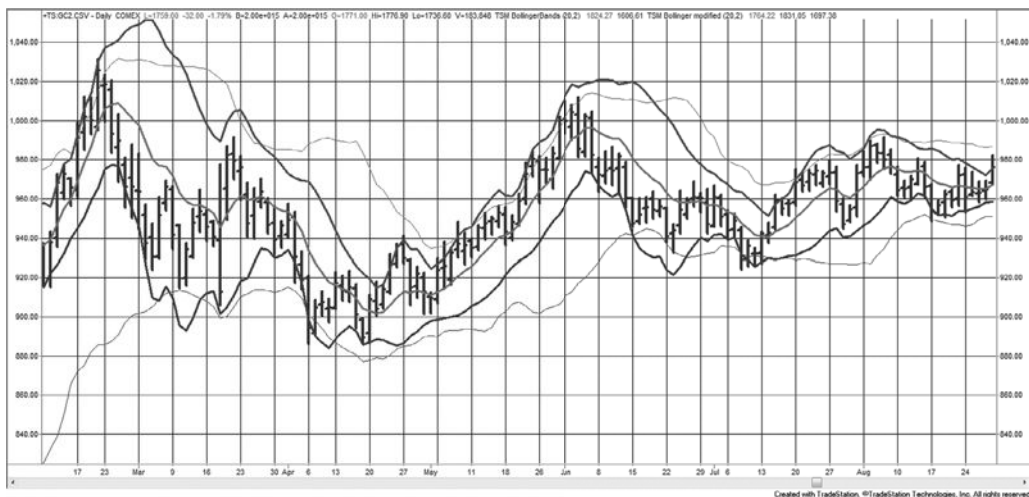
FIGURE 8.8 Modified Bollinger bands shown with original bands (lighter lines), applied to gold futures, February–August 2009.

$$d_t = \frac{(2-\alpha) \times m_t - u_t}{1-\alpha}$$

$$BU_t = D_t + f \times d_t$$

$$BL_t = D_t - f \times d_t$$

where f is the multiplier for the width of the band, suggested at 2.5 compared to Bollinger's 2.0. Figure 8.8 shows the modified Bollinger bands along with the original (lighter lines) for gold futures during the first part of 2009. While the new bands do not remove the bulge, they are faster to correct and more uniform in the way they envelop prices. Programs to calculate and display the original and modified bands are *TSM Bollinger bands* and *TSM Bollinger Modified*, available on the Companion Website.

Rules for Using Bands

Regardless of the type of band that is constructed, rules for using bands to generate trading signals are limited. The first decision to be made is whether the trading strategy is one that is always in the market (a reversal strategy), changing from long to short and back again as the bands are penetrated. If so, the following rules apply:

- *Buy* (close out shorts and go long) when the prices close above the upper band.
- *Sell short* (close out longs and go short) when the prices close below the lower band.

This technique is always in the market with a maximum risk (without execution costs) equal to the width of the band, which changes each day (see Figure 8.9). Alternately, you may prefer to exit from each trade when prices move into the zone between the bands or when prices cross the original trendline.
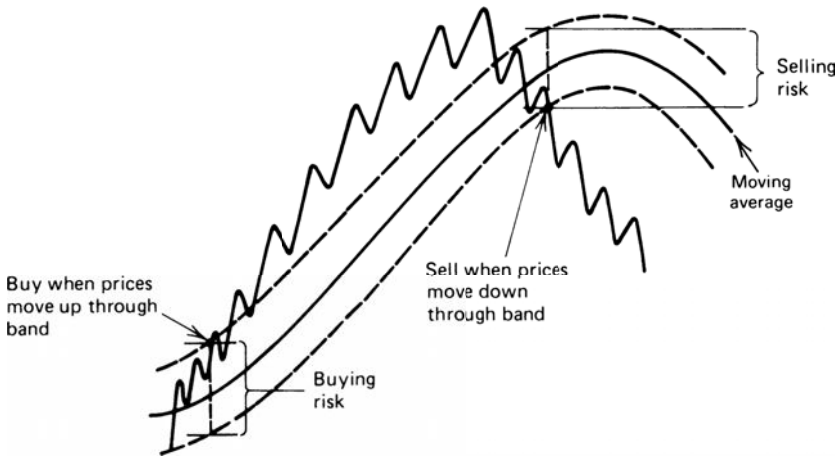
FIGURE 8.9   Simple reversal rules for using bands.

- *Buy* (go long) when prices close above the upper band. Close out longs when prices reverse and close below the moving average value (the center of the band).
- *Sell short* when prices close below the lower band. Cover your shorts when prices close above the moving average value.

   The band is then used to enter into new long or short trades, and the actual trendline at the center of the band is used for liquidation. If prices are not strong enough to penetrate the opposite band on the close of the same day, the trade is closed out but not reversed. The next day, penetration of either the upper or lower band will signal a new long or short trade, respectively.

   This technique allows a trade to be reentered in the same direction in the event of a false trend change. If a pullback occurs after a close-out while no position is being held (as shown in Figure 8.10), an entry at a later date might be at a better price. It also reduces
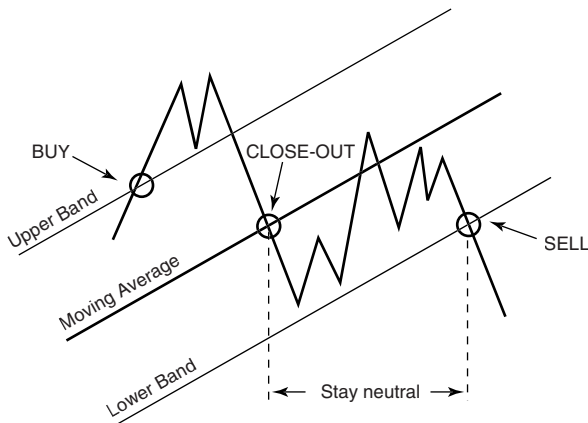
FIGURE 8.10   Basic rules for using bands.

the order size by 50%, which is likely to improve the execution price and add liquidity for large traders. The disadvantage is when the price changes direction and moves so fast that both the close-out and the new signal occur on the same day. Reversing the position immediately would be better in a fast market.

The high and low of the day may also be used as penetration criteria. Again using the outer bands for entry and the moving average for exit, apply the following rules:

- *Buy* when the high of the day penetrates the upper band, and close out longs when the low of the day penetrates the moving average.
- *Sell short* when the low of the day penetrates the lower band, and cover shorts when the high penetrates the moving average.

Using the trendline as an exit, risk is limited to half of the full band width. If the bands are narrow, there is a greater chance that an entry on an intraday high might also see an exit below the trendline on the close of the same day.

Timing the Order

The type of execution order placed when following a system will affect its results over the long term. The use of a moving average band identifies a change of trend when a breakout occurs. Buying at the point of the upside breakout or selling during a downside breakout often results in poor entry prices, and has been known to place the trader in a new trend at the point where prices are ready for a technical correction. In an attempt to overcome these problems, a number of rules can be used:

- *Buy* (or *sell*) on the close after an entry (intraday) signal has been indicated.
- *Buy* (or *sell*) on the next day's open following a signal.
- *Buy* (or *sell*) with a delay of 1, 2, or 3 days after the signal.
- *Buy* (or *sell*) after a price retracement of 50% (or some other value) following a signal.
- *Buy* (or *sell*) when prices move to within a specified risk level relative to a reversal or exit point.

The object is to enter a new position and see an immediate profit, or reduce risk. Some of these rules can be categorized as timing and others as risk management. If intraday prices are used to signal new entries and exits, a rule may be added that states:

Only one order can be executed in one day; either the liquidation of a current position or an entry into a new position.

While better entry points will improve overall performance, an entry rule that is contingent on price action, such as a pullback, risks the possibility of not entering at all. A contingent order that is missed is guaranteed to be a profit. It might be better to combine the entry order, for example,

Buy (or sell) after prices reverse by 0.50 × ATR or enter on the next close.

Once you have decided on a timing rule, you must test it carefully. The perception of improvement does not always live up to expectations. In tests on trend-following systems conducted over many years, positions calculated on the close but delayed until the next open improved execution prices about 75% of the time but resulted in smaller overall total profits. Why? Fast breakouts that never retrace result in missed trades. Therefore, while three out of four executions returned a better price by a small amount, those improvements were often offset by the profitable breakouts that were never entered.

The Compromise between Reliability and Delay

As with most trading techniques, the benefits of one approach can also have negative factors. The use of a band around a trendline improves the reliability of the trading signal and reduces the total number of signals. The wider the band, the fewer signals. Both of these characteristics are significant benefits. But wider bands mean delayed entries; therefore, you cannot capture as much of the trend, and the average profits will be smaller. If the bands are too wide, then the average profits can decline to zero. The use of wider bands also means greater risk on each trade. It will be necessary to trade smaller positions or capitalize the account with a larger investment.

These are serious choices that must be made with every trading program. Although there are classic solutions to this problem discussed in Chapters 23 and 24, traders must choose the methods that complement their risk preference.

Bollinger on Bollinger Bands

While most trading strategies buy when there is an upwards penetration of the top band and sell when prices move below the lower band, the use of Bollinger bands is usually mean-reverting, or counter to the price direction. However, this can be unnecessarily risky, especially when prices are volatile. Bollinger recommends confirming a downside penetration using other indicators, primarily those based on volume and market breadth. If prices are moving lower but volume is not increasing and negative breadth is not confirming the downward move, then a buy signal is realistic.[4]

Bollinger uses the concept that volatility is cyclic, but without a regular period. He sees very low volatility as a forecast for high volatility and very high volatility forecasting low volatility, similar to the way traders use the CME Volatility Index (VIX). Based on this, a major price rally with dramatically higher volatility, that expands the bandwidth to extremes, should be sold when the bandwidth begins to narrow. This only applies to upwards price moves.

---

[4]John Bollinger, "John Bollinger of Bollinger Bands Fame," *Technical Analysis of Stocks & Commodities* (May 2002).

Combining Bollinger with Other Indicators

Williams[5] suggests that a number of indicators can be combined to capture volatile moves after a price contraction, using:

1. A standard 20-day, 2-standard deviation Bollinger band
2. A 20-day Keltner Channel
3. A 21-day Chaikin Oscillator to monitor the flow of funds

To enter a long position, the following conditions must be satisfied:

- The Bollinger bands constrict to inside the Keltner Channel while the Chaikin Oscillator is below zero.
- The Chaikin Oscillator crosses above the zero line.

For shorts,

- The Bollinger bands constrict to inside the Keltner Channel while the Chaikin Oscillator is above zero.
- The Chaikin Oscillator crosses below the zero line.

APPLICATIONS OF A SINGLE TREND

For any trend technique, the selection of the calculation period—the interval over which you will define the trend—is the most important decision in the ultimate success of the trading system. Entry rules and timing improve performance but are considered refinements. The calculation period determines the frequency of trading and the nature of the underlying trend that will be targeted. Deciding the calculation period is more important than the method of identifying the trend. You can be profitable using a simple moving average, regression, breakout, or any other technique—if you can settle on the right time interval.

The previous sections have used examples of calculation periods without any claim that one time interval was better than another. We have discussed that the long-term trend mimics government policy of interest rates or economic growth; therefore, there is good reason to choose a longer calculation period. We also saw that the trends were clearer when looking at a weekly chart rather than at a daily, and it was not clear that an intraday chart had any persistent trends. But for most traders, the risk of using this long time frame is unacceptable; they prefer smaller profits and smaller losses associated with

---

[5]Billy Williams, "Biting Off Profits with the Rattlesnake Breakout Method," *Futures* (October 2010).

faster trading. There is now strategy development software that makes it easy to test a range of calculation periods to find the one that performed best in the past. This technique is called optimization and is discussed in Chapter 21. But the power of the computer is not always as good as simple human reasoning and common sense. The computer is best for validating an idea, not for discovering one.

Before the computer, analysts struggled with the same problem of finding the best calculation period. At first, the trend period was based on multiples of calendar periods, such as a week or a month, expressed as trading days. When these techniques were limited to a small group of analysts, these approaches were very successful. Many traders still subscribe to the idea that certain time intervals have value. The most popular calculation periods have been: 3 days, the expected duration of a short price move; 5 days, a trading week; 20 to 23 days, a trading month; 63 days, a calendar quarter; and of course, 252 days, a calendar year. Implied volatility calculations traditionally use 20 days. It is not clear where the 200-day moving average, used for stocks, came from.

More recently, a class of adaptive trends has appeared. These techniques attempt to change the speed of the trend based on a characteristic of price movement, such as volatility or noise. These techniques are another alternative to optimization and are discussed in Chapter 17.

The following sections include classic examples of well-known systems that use one trend, as well as a comparison of trading performance of the most popular single-trend techniques over a broad range of calculation periods.

A Simple *Momentum* System

In Chapter 7 the *n-day momentum* was defined as the change in price over *n* days. It's not actually "momentum" but that is the term commonly used by the industry. The simplest trend system is the one that buys when the *n*-day change is positive and sells when the *n*-day change is negative. For large values of *n*, the results will be surprisingly similar to a simple moving average system; therefore we will not give examples here. Keep in mind that momentum can be very effective even as it is very simple.

A Step-Weighted Moving Average

In 1972, Robert Joel Taylor published the "Major Price Trend Directional Indicator" (MPTDI), which was reprinted in summary form in the September 1973 *Commodities Magazine* (now *Futures*). The system was promoted and implemented through Enterex Commodities in Dallas and was tested in 1972 on historical data provided by Dunn and Hargitt Financial Services in West Lafayette, Indiana. It was one of the few well-defined published systems and served as the basis for much experimentation for current technicians and aspiring analysts.

MPTDI is a moving average with a band. Its unique feature is that the calculation period and band width change based on price volatility, the current trading range. Because

TABLE 8.6   MPTDI Variables for Gold*

Average Trading Range	Number of Days in Calculation	Weighting-Factors Progression	Entry Signal Penetration	Approximate Stop-Loss Point
50–150	2–5 days	TYPE A	100 pts	150 pts
150–250	20 days	TYPE B	200 pts	300 pts
250–350	15 days	TYPE C	250 pts	350 pts
350–450	10 days	TYPE D	350 pts	450 pts
450+	5 days	TYPE E	450 pts	550 pts

*100 points = $1 per ounce.

the method has distinct trading range thresholds (called steps), the method is called a *step-weighted moving average*. It is unique in its complete dependence on incremental values for all aspects of the system: the moving average, entry, and stop-loss points. For example, Table 8.6 shows what conditions might be assigned to gold.

If gold were trading in an average range of 250 to 350 points each day ($2.50 to $3.50 per ounce, but remember this was 1972), the weighting factor for the moving average would be TYPE C, indicating medium volatility (TYPE A is lowest). Using TYPE C with a 15-day moving average, the most recent 5 days are given the weight 3, the next 5 days 2, and the last 5 days are weighted by 1. The entry signals use the corresponding penetration of 250 points above the moving average for a buy and 250 below for a sell. The intraday highs or lows are used to trigger the entry based on values calculated after the close of trading on the prior day. A stop-loss is fixed at the time of entry equal to the value on the same line as the selected volatility. The penetration of the stop-loss will cause the liquidation of the current trade. A new signal in the reverse direction will serve as both the exit for the current trade and the entry for a new trade.

There is a lot to say in favor of the principles of MPTDI. It is individualized with respect to markets and self-adjusting to changing volatility. The stop-loss serves to limit the initial risk of the trade and allow the coordination of a money management approach. The fixed risk differs from moving averages using standard bands because a moving average and its band can back away from system entry points if there is a gradual reversal of the price trend. But there are some rough edges to the system. The incremental ranges for volatility, entry points, and stops seem to be a crude measure. Even if they are accurate in the center of the range, they must get less accurate at the extremes where volatility causes an abrupt change in parameter values when it moves from one range to another.

MPTDI sets the groundwork for a smoother, more adaptive process. Before such a process can be developed, however, it is necessary to study price movement at discrete levels, such as those shown in MPTDI. From discrete relationships it is possible to generalize a continuous relationship. These methods are covered in Chapter 17.

The Volatility System

Another method that includes volatility and is computationally simple is the *Volatility System*.[6] Signals are generated when a price change is accompanied by an unusually large move relative to average volatility. If the average volatility measured over n days is

$$V_t = \frac{1}{n} \sum_{i=t-n+1}^{t} TR_i$$

where TR_i is the *true range* on day i and V_t is the called the *average true range* on day t (defined in Chapter 5).

Trading rules are given as

- *Sell* if the close drops by more than $k \times V(n)_{t-1}$ from the previous close.
- *Buy* if the close rises by more than $k \times V(n)_{t-1}$ from the previous close.

The value of k is generally about 3.0. Note that the current price change is always compared to the previous volatility calculation.

The 10-Day Moving Average Rule

The most basic application of a moving average system was proposed by Keltner in his 1960 publication, *How to Make Money in Commodities*. Of three mechanical systems presented by Keltner, his choice of a moving average was based on performance and experience. The system itself is quite simple, a 10-day moving average using the average of the daily high, low, and closing prices, with a band on each side formed from the 10-day moving average of the high-low range (similar to a 10-day average true range). A buy signal occurs on penetration of the upper band and a sell signal when the lower band is broken; positions are always reversed.

The *10-Day Moving Average Rule* is basic, but it does apply the fundamental volatility principle by using the high-low range as a band, and serves as an early example of moving averages. Keltner expresses his preference for this particular technique because of its identification of minor rather than medium- or long-term trends, and there are some performance figures that substantiate his conclusion. As an experienced trader, he prefers the speed of the 10-day moving average, which follows the market prices with more reasonable risk than slower methods. A side benefit to the selection is that the usual division required by a moving average calculation can be substituted by a simple shift of the decimal place; in an era before the pocket calculator, who knows how much impact that convenience had on Keltner's choice.

The history of prices now shows us that price movement was much smoother up to the end of the 1970s and has been getting noisier ever since. A 10-day moving average, supplemented by a volatility band, was truly the state-of-the-art technology. While the

---

[6]Richard Bookstaber, *The Complete Investment Book* (Glenview, IL: Scott, Foresman, 1984), 231.

shorter calculation periods are not generally successful for current price moves, the use of volatility to create bands has held up well over time.

TRIX, Triple Exponential Smoothing

A triple exponential smoothing technique was first described by Hutson as another approach to trend following[7] and over the years has gained in popularity. Called *TRIX*, it first takes the log of the price to account for growth and then applies an exponential smoothing three times using the same smoothing constant. A buy signal is generated when the triple-smoothed trendline rises for two consecutive days; a sell signal followed a 2-day decline in the trendline. The exponential smoothing process usually starts by setting the initial trend value $E1_0 = p_0$, but in this case $E1_0 = \ln p_0$. The rest of the process is

$$E1_t = E1_{t-1} + s \times (\ln p_t - E1_{t-1})$$
$$E2_t = E2_{t-1} + s \times (E1_t - E2_{t-1})$$
$$E3_t = E3_{t-1} + s \times (E2_t - E1_{t-1})$$
$$\text{TRIX} = (E3_t - E3_{t-1}) \times 10000$$

This original approach has seen some variations over the years. The most significant is not using the log of prices, but changing the final step to a percentage change. The percentage change at the end speeds up the process. In all cases, the smoothing constant should represent a short time period, less than 20 days, but recommended as 6 days. The number of days is converted to a smoothing constant using $s = 2/(n + 1)$. The alternative calculation is

$$E1_t = E1_{t-1} + s \times (p_t - E1_{t-1})$$
$$E2_t = E2_{t-1} + s \times (E1_t - E2_{t-1})$$
$$E3_t = E3_{t-1} + s \times (E2_t - E1_{t-1})$$
$$TRIX_t = (E3_t - E3_{t-1})/E_{t-1}$$

A *signal line* is created by taking the 3-day moving average of the most recent TRIX values. A buy occurs when TRIX crosses above the signal line and a sell when it crosses below the signal line. Using a signal line is a technique that will be seen with other momentum indicators.

A 9-day TRIX is shown in the lower part of Figure 8.11 corresponding to the price of the EURUSD (euro currency). The final step that takes the difference between the current and previous TRIX value shifts the indicator so that it does not have the lag that would be expected, yet it is still smooth. The effect of the weighting on price data caused by double and triple smoothing was discussed in Chapter 7. Readers that are interested in similar methods should refer to Blau's *True Strength Index* and *True Directional Movement*.

---

[7]Jack K. Hutson, "Good TRIX," *Technical Analysis of Stocks & Commodities* (July 1983).
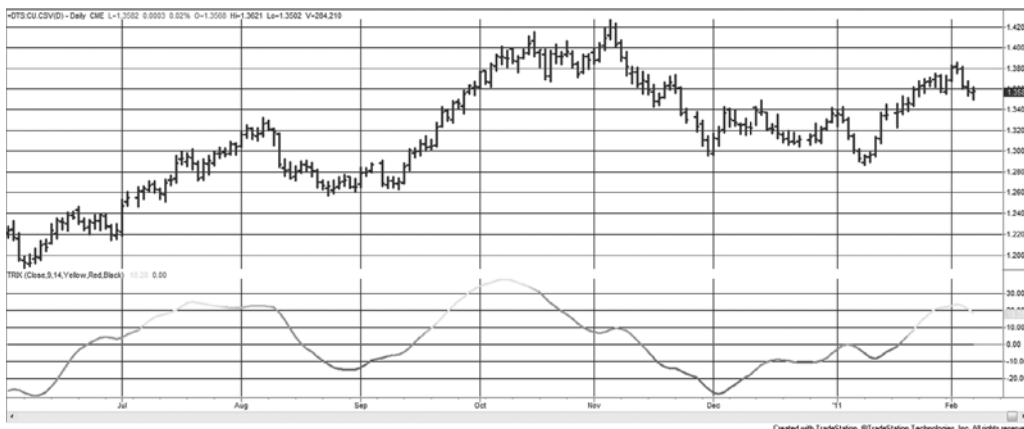
FIGURE 8.11 A 9-day TRIX based on euro futures shows that a triple smoothing does not create the lag that would be expected.

Raschke's First Cross

Although we usually trade a trend from the beginning to end based on some smoothing method, Linda Bradford Raschke has shown that a selected piece of the trend move can be a very reliable trade. It is also necessary to introduce the idea of momentum, which will be the subject of the next chapter. Her *First Cross* system enters a trend trade on the first pullback after an initial trend signal based on a faster momentum indicator. *Momentum* will be used here as *the difference between two trends*. To create this strategy we need

```
slowMA   Slow moving average
fastMA   Fast moving average
osc      Moving average oscillator
trend    Moving average of oscillator
```

The *oscillator* is the difference between the fast- and slow-moving averages

$$osc = fastMA - slowMA$$

A buy or sell entry signal is a 3-step process (comments on right):

B1.	$osc_{t-1} > trend_t$ and $osc_t <= trend_t$	*osc* crosses trend moving down
B2.	$low_t > low_{t-1}$	low of the current bar is rising
B3.	*Buy*	*buy* signal
S1.	$osc_{t-1} < trend_t$ and $osc_t <= trend_t$	*osc* crosses trend moving up
S2.	$high_t < high_{t-1}$	high of the current bar is declining
S3.	*Sell*	*sell short* signal

Although Raschke does not discuss exits, a likely exit rule for the First Cross system should maintain its short-term nature. You might use this method to time an entry into

a longer-term trend using the first pullback. Or, you might want to trade only the short period isolated by the first pullback until this initial impulse is exhausted. If you choose the second method, you might try the following sequence for an exit condition:

X1. $close_{t-1} > close_{t-2}$ previous close turns up (confirms trend)
X2. $close_t < close_{t-1}$ the current close turns down
X3. *Exit long*

Raschke's idea is an excellent example of selectivity. First, you recognize that the beginning of a trend is unique event. As traders recognize a change of direction, the move strengthens. Because most systems lag the market, they are often too late in capturing initial profits. As an alternative, this technique waits for the first move to be exhausted and then enters in anticipation of another surge as the new trend reasserts itself. Once these early moves are over, the general trend move may not be as easy to work with, and you may find yourself trying to enter as the trend comes to an end.

COMPARISON OF MAJOR TREND SYSTEMS

Trend strategies dominate the world of algorithm trading and managed futures in particular. But which method is the best? As we will see later in this book, there are many rules that can be added to a basic trend strategy including stop-losses, profit-taking, and entry timing, that change both the returns and the risk profile. There are cases where an underlying losing strategy can be turned profitable by risk management or clever timing rules; however, it is always best to start with a sound trend-following method that has the risk characteristics most acceptable to you.

This chapter will not draw conclusions about which trending method is best. It may be that there is no best strategy, only trade-offs between risk and reward, fast or slow. By testing a small sample of markets for the same time period and a selection of calculation periods, we can understand how the major trending methods differ. The most popular approaches, two event-driven (discussed in Chapter 5) and four time-driven are:

1. *M, N*-day momentum
2. *MA*, Simple moving average
3. *EXP*, Exponential smoothing
4. *NDB, N*-day breakout
5. *SWG*, Swing breakout
6. *LRS*, Linear regression slope

The markets used will be IBM, Ford, and Bank of America representing the equities, and Eurodollar interest rates, the S&P 500, the euro currency (EURUSD), and crude oil

futures markets. The data for equities has been adjusted for splits, and the futures are continuous, back-adjusted. Neither of these data adjustments affects the trend calculations or trading signals. The time period will be the most recent 20 years. For the time-based systems, the calculation periods will begin at 5 and test every 5-day interval up to 160. If it wasn't so easy to do this in a test platform, we could get a representative sample by starting at 5 and doubling in order to keep the percentage change the same, such as 5, 10, 20, 40, 80, and 160. That would also remove the distortion if we were to look at average results, because the longer calculation periods, which are very similar, would outweigh the shorter periods. For the swing method, a varying percentage swing size will be used, starting at 0.25% and doubling to 4%. The data used was from CSI and the testing was done using TradeStation. All tests begin trading on the same date, even though the 5-day test needs less data to start the calculations than the 160-day test.

Trading Rules

To see the characteristics of each system, the trading rules will be as simple as possible. Only the basic buy and sell signals will be used (where *sell* is both exiting longs and selling short). All six systems are always in the market. That is, once they enter a long or short position, they always reverse on a new signal, going from long to short, or short to long. There are no stop-losses or other risk controls; therefore, each system will show its own, natural risk profile. All entries and exits are done at the closing price. A $20 transaction charge was applied to each futures trade, and $0.01 per share for each stock trade to cover commissions and slippage. Without a transaction charge, the very fast trading systems will show much better results than they can achieve in real trading. The following is a brief description of the type of system, calculation method, and trading rules. Note that, for MA, EXP, and LRS, the trading signal is based on the direction of the trendline.

1. *M, N*-day momentum
   a. *Buy* when $close_t > close_{t-n}$
   b. *Sell* when $close_t < close_{t-n}$
2. *MA*, Simple moving average
   a. *Buy* when $MA_t > MA_{t-1}$
   b. *Sell* when $MA_t < MA_{t-1}$
3. *EXP*, Exponential smoothing.
   a. *Buy* when $Exp_t > Exp_{t-1}$
   b. *Sell* when $Exp_t < Exp_{t-1}$
4. *NDB, N*-day breakout
   a. *Buy* when $high_t > highest(high, t-1, n)$ *and* $close_t > close_{t-1}$
   b. *Sell* when $low_t < lowest(low, t-1, n)$ *and* $close_t < close_{t-1}$

5. *SWG*, Swing breakout

   a. *Buy* when *the current swing high > previous swing high*

   b. *Sell* when *the current swing low < previous swing low*

6. *LRS*, Linear regression slope

   a. *Buy* when *Slope(close,t,n) > 0*

   b. *Sell* when *Slope(close,t,n) < 0*

In the rules described above, the functions *highest* and *slope* use the parameters (*price*, *current day*, *calculation period*), then *highest*(*high*, *t*–1, *n*) will return the highest high for the *n* days ending on the previous day, *t*–1.

Spreadsheet Example

A spreadsheet is an easy way of seeing the returns of all but the swing method. The function *OFFSET* allows the calculation period (located in F3) to be changed, resulting in all calculations and returns changing. It is a simple way of allowing different calculation periods to be tested. The calculations for the other five can be done in a single column, using the following setup and instructions:

1. Column A is the date.

2. Columns B, C, D, and E are the open, high, low, and closing prices.

3. F3 has the calculation period that will be used for all five strategies.

4. Column F calculates the momentum as = E163 – OFFSET(E163, –$F$3,0). Calculations begin in row 163 because there will be a maximum of 160 days allowed.

5. Column G is the moving average = AVERAGE(E163: OFFSET(E163, –$F$3,0)).

6. Column H is the exponential smoothing, = H162 + $H$3*(E163 – H162). Cell $H$3 = 2/(F2 + 1), the standard conversion from days to smoothing constant.

7. Column I is the regression slope = SLOPE(E163:OFFSET(E163, –$F$3,0),A163: OFFSET(A163, –$F$3,0)).

8. Column J records if the most recent breakout is up (+1) or down (–1) = IF(E163 > MAX(E162: OFFSET(E162, –$F$3,0)),1,IF(E163 < MIN(E162: OFFSET(E162, –$F$3,0)), –1,"")).

The next five columns, K–O, show the continuous trend direction (+1 or –1) based on the calculations in F–J. Once there is an initial direction, the cells are either +1 or –1. The breakout strategy, column J, takes 77 days before the first trend can be identified.

The last five columns give the cumulative profit or loss in points; that is, there is no conversion to dollars. For the Eurodollar interest rates, the futures market conversion is $2500 per big point, making a move from 97.00 to 98.00 worth $2,500. No commissions are used, although those costs could significantly affect the comparison. Results for Eurodollars, the S&P, and IBM can be found in three spreadsheets, *TSM Trend Systems Comparisons ED/SP/IBM*, available on the Companion Website.
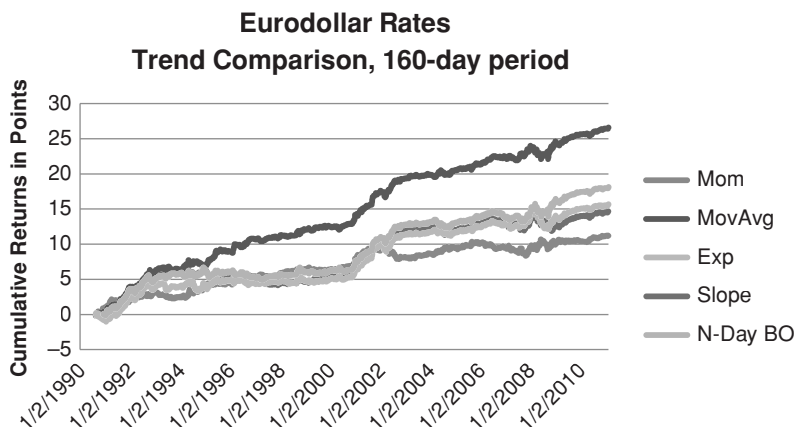
FIGURE 8.12 Cumulative returns of Eurodollar futures for the five spreadsheet strategies.

Spreadsheet Results of Eurodollars, the S&P, and IBM

Figure 8.12 shows the cumulative returns of the Eurodollar futures using the 160-day calculation period, the longest to be tested. Over the 20-year test period, the moving average performed the best, with the n-day breakout second and the n-day momentum at the bottom. The best returned about 26 points and the worst about 11 points. In dollars, that would range from a simple return of $3250 to $1375 per year. For an investment of three times the typical initial margin of $1500, that gives a return of 72% and 30% per annum. Of course, there is no commission and no slippage used, and Eurodollars have the most trending characteristics, so this should not be taken as representative of all performance. A more complete review will follow in the next section.

By observing the performance pattern of the five systems, we can draw some interesting and valuable conclusions:

- Long-term trend following can be profitable.
- All basic trending strategies are profitable if the market trends.
- When adding other features to a system, it needs to be proved that those features improve the results, because the simple approach seems to be very good.

S&P

The S&P is a much noisier market than Eurodollars, which shows in the Figure 8.13 performance. Again, the moving average leads with the momentum showing the only outright loss over the test period. The jump in the moving average performance was the result of entering a short sale on April 24, 2007, and holding it until October 16, 2008. While this shows the advantage of the long-term trend, most traders would find it difficult to watch the weekly and monthly fluctuations in their equity over this holding period.

S&P *e*mini
Trend Comparison, 160-day period
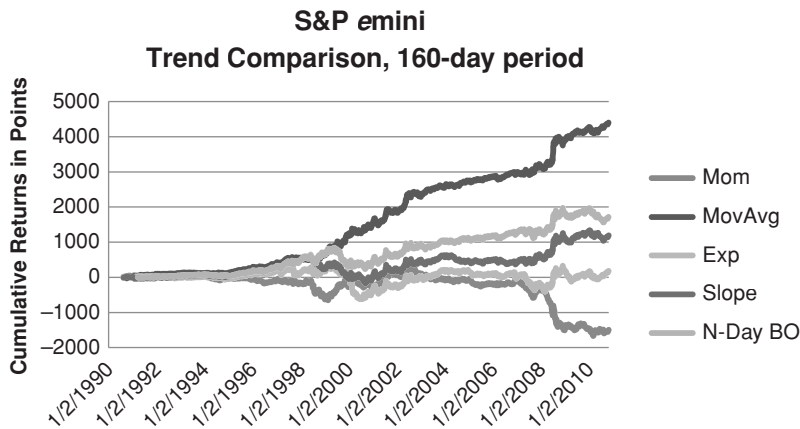
FIGURE 8.13 Comparative trend results for the S&P *e*mini futures.

IBM

The last example in Figure 8.14 is IBM, which should bear some relationship to the S&P, given that it is a high-cap stock and must also be a relatively noisy market. In fact, these results show that all returns are worse than the S&P index, with the moving average the only profitable method. It is likely that the index smoothes out the erratic moves of individual stocks. But a brief look at the results using only one calculation period doesn't show the whole picture.

Comparative Results

For comparison, four futures markets will be used, Eurodollar interest rates (ED), euro currency (CU), the S&P (SP), and crude oil (CL), and three stocks, IBM, Bank of America

IBM
Trend Comparison, 160-day period
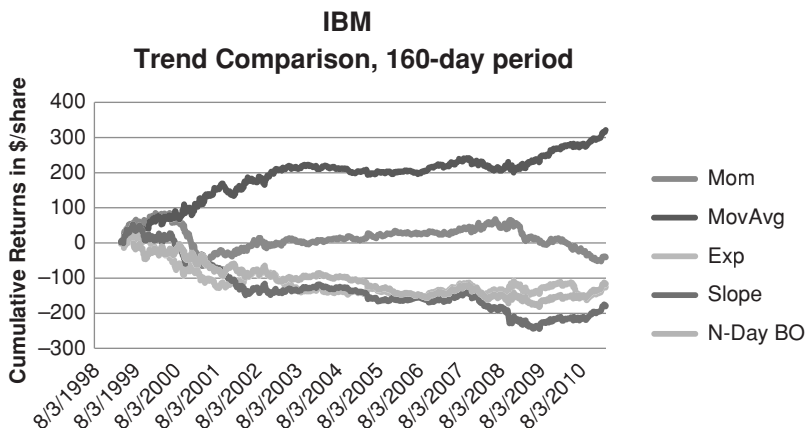
FIGURE 8.14 Comparative trend results for IBM from August 1998 through February 2011.

Moving Average Net Profits—Stocks
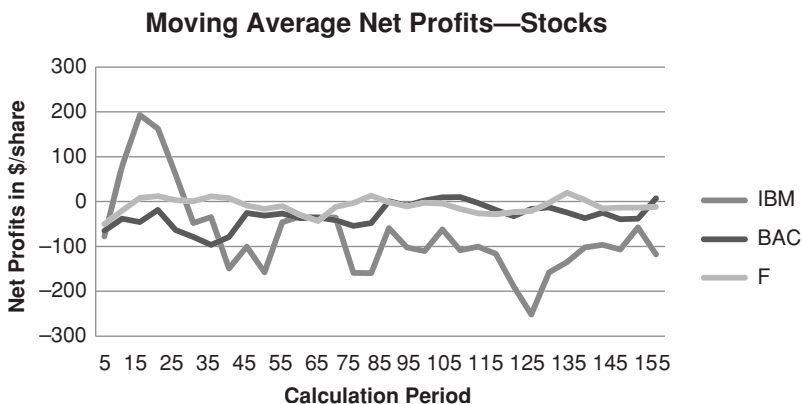
FIGURE 8.15a Moving average net profits for futures markets, 20 years ending February 2011.

Moving Average Net Profits—Futures
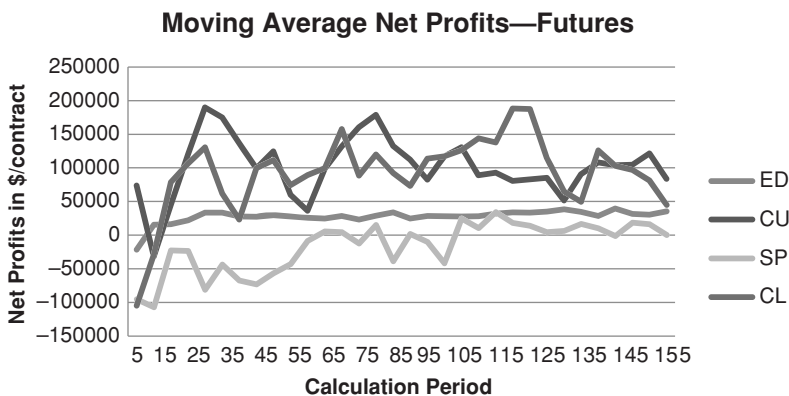
FIGURE 8.15b Moving average net profits, 12 years ending February 2011.

(BAC), and Ford (F). The test results will be shown as charts because visualization makes it easier to understand.

The first comparison is *net profit* because everyone wants to know how much each system and each market will return. Looking only at the moving average tests (Figure 8.15b), Eurodollars, the euro, and crude oil are profitable to different degrees across most of the calculation periods, but tend to lose in the very short-term. The S&P generates mostly losses but more in the short-term.

The three stocks show losses for nearly all periods (see Figure 8.15a). In terms of consistency, the Eurodollar interest rates stand out as very stable from 25 to 160 days compared to the other six tests. This can be attributed to its relatively low noise compared to all other markets, as discussed in Chapter 2. Because equity markets are at the other end of the noise spectrum, trends are more likely to fail. Is this conclusion the same over all trending methods, or is one better than another for specific markets?

Different Trend Philosophies

The moving average, exponential, and linear regression are traditional time-series processes, that is, the price of one day is a small part of the total calculation. This can make it very difficult for a single large price jump to cause a change in the direction of the trend and, consequently, the trading signal. It also introduces a lag. For very fast trends, each price will have more importance than for slower trends.

The n-day breakout and swing breakout systems are triggered on new highs and lows; therefore, a single big price move often causes a new signal. Breakout systems can be considered event-driven, while a moving average system requires a trend to evolve. Because an event-driven system gives a signal at the moment prices make new highs or lows, there is no lag.

Different Risk Profiles

The biggest difference in the trend methods is in their risk profiles. Even when the returns are similar, the risk of the moving average technique is remarkably different from the breakout methods. While the information ratio (annualized return divided by annualized risk) is the best way to distinguish one system from another, the *profit factor* (gross profits divided by gross losses) was the only measurement available for these tests; however, the results will still give a good idea of the relative difference in performance. Figure 8.16a shows the profit factors for the Eurodollar interest rate futures for 5 of the 6 methods, all based on the same calculation periods. Data is for 20 years through 2011. Figure 8.16b shows the profit factors for the swing method, which varies the percentage swing size.

Even though we recognize the Eurodollar as a highly trending market, the similarity in the profit factors is remarkable. The n-day breakout seems to be slightly better than the other methods, but all of them are clustered together, improving at the same rate as the calculation period increases. This brings up the question "Does it really matter which trend method we use?" Are trend profits the result of a clever formula, or is it the market that controls the result?

The swing method has a problem for interest rates because tests are based on varying the percentage swing. But rates are quoted in prices, not yields, so volatility goes down as prices go up. Figure 8.16b shows that, using the range of percentage swings that work for other markets, the results using prices had no trades when the swing was above 1.375%. Using the yield, by subtracting the price from 100, gave trades everywhere but had poor returns. This example is included here to point out that any analysis of interest rate futures prices that use percentages will be a problem.

Based on the profit factor, five of the systems look the same for the Eurodollar interest rates. Other markets are not as consistent. The euro currency, Figures 8.17a and b, show that the n-day breakout is quite different for long calculation periods. The rules of the breakout only change positions on new highs and lows, and a calculation period over 125 days is equal to 6 months. Then a new long is set on a high above the highest price of the past 6 months and not exiting until a new low of the past 6 months. That can create a lot of risk. The same is true, to some degree, of the swing method. A swing of 5% in crude oil at $30/bbl is $1.50, but at $100/bbl it's $5.00. Even though a percentage swing is somewhat self-adjusting in sensitivity, the risk continues to increase (see Figures 8.18a and b).

Eurodollar Profit Factors
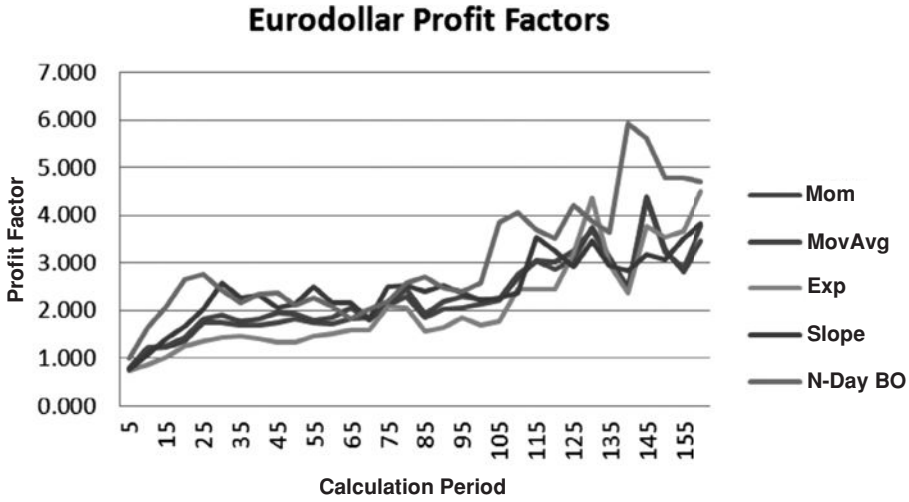
FIGURE 8.16a Profit factors for five of the systems show that returns improve as the calculation period increases.

Eurodollar Swing Profit Factors
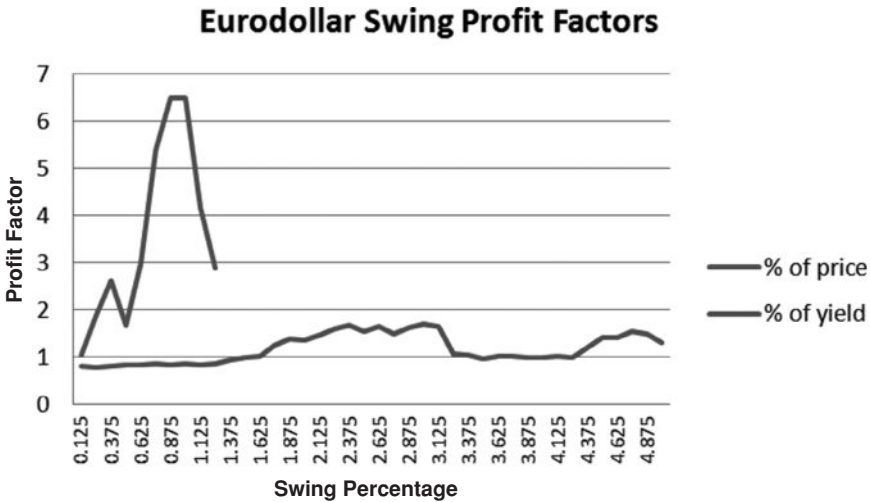
FIGURE 8.16b Profit factors for the swing method for interest rate markets use a percentage of either price or yield.

The S&P, along with the euro and crude oil, shows similar profit factors for all trend methods from the fastest calculation periods through about 85 days (see Figures 8.19a and b). After that returns begin to diverge and become more erratic. The n-day breakout is consistently the best for longer periods, but also highly variable.

But that's still not the whole story. The frequency of trades, the percentage of profitable trades, and the average profit highlight the differences in the methods. Using only the euro currency as an example because it seems most representative, Figure 8.20a shows the number of trades for each calculation period, Figure 8.20b the percentage of profitable
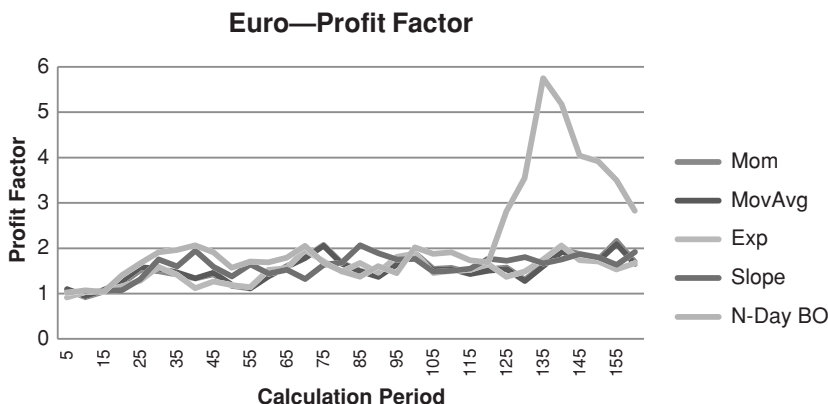
Euro—Profit Factor

FIGURE 8.17a Euro currency futures profit factors for 5 trend systems over a range of calculation periods for 20 years through 2011.
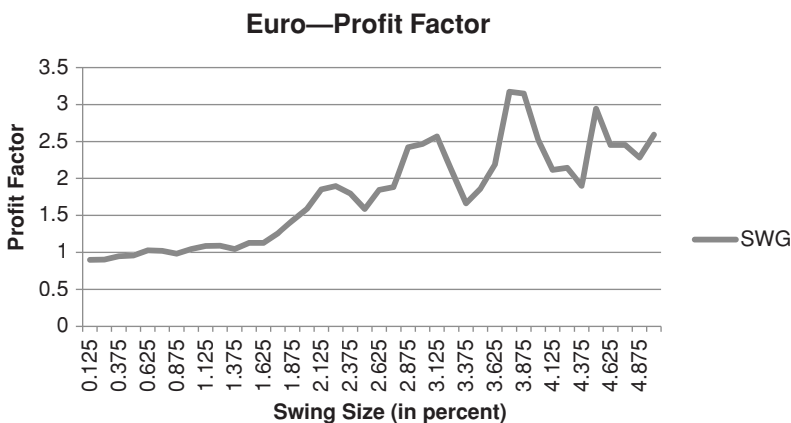
Euro—Profit Factor

FIGURE 8.17b Euro currency profit factors for swing method, continuous futures, 20 years through 2011.

trades, and Figure 8.20c the average profit per trade. Again, the breakout system stands out as having less than half the number of trades of the other trend methods. If we use the moving average as an example, prices can move up and down at the time of the trend change causing frequent buy and sell signals until a new trend is decided. These losses are small, but there can be many of them. The risk of a trade in the breakout system is the difference between the n-day high and n-day low, which can be very large but not likely to be reached often. Then we would expect fewer trades and larger losses for the breakout system.

The percentage of profitable trades, also called *reliability*, varies considerably for each method with the moving average and momentum results showing the lowest reliability at about 34% (results are almost identical) and the breakout system at the top with just under 50%. The slope and exponential smoothing are both about 43%. This chart tells

Crude Oil Profit Factors
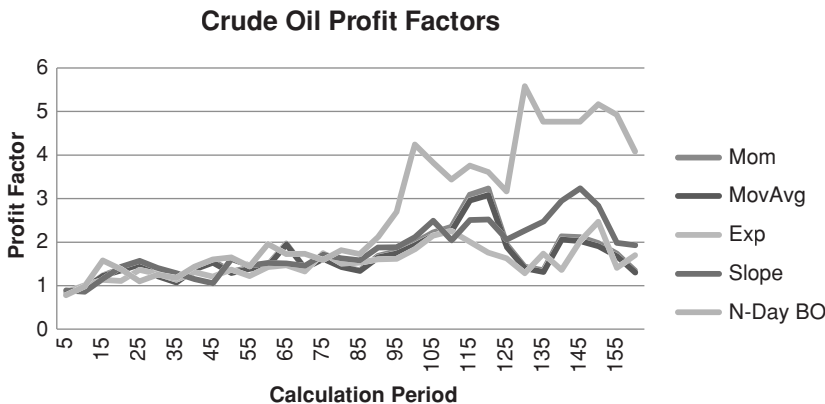
FIGURE 8.18a Crude oil futures, profit factors for 5 trending systems, 20 years through 2011

Crude Oil Swing Profit Factor
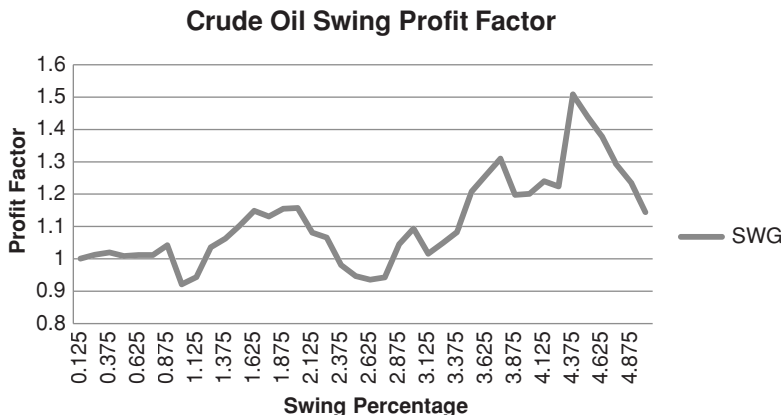
FIGURE 8.18b Crude oil futures, profit factors for swing method, 20 years through 2011.

us that, although the breakout system holds trades longer and has greater risk, it is more likely to end with a profit.

The last chart, average profit per trade, reflects the holding period of the methods. The longer you hold a trade, the larger the average profit, although the average risk will also get larger. The moving average, exponential smoothing, and momentum are all similar and almost indistinguishable on the chart. The regression slope has much larger profits. But the breakout stands alone as having very large profits at all calculation periods.

The initial results that we saw, using the simplified spreadsheet and a calculation period of 160 days, are not as clear as when we see more of the system profile. The moving average is still a consistent top performer, but the breakout often has higher returns at the same time it has higher risk. One thing is clear—all of these trend systems are profitable in the long term.
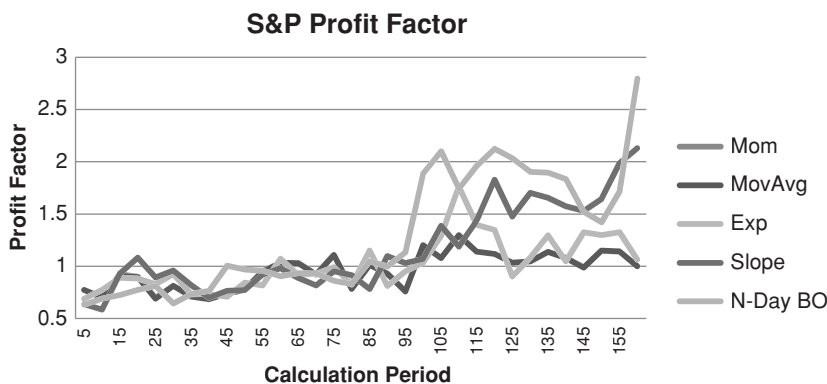
FIGURE 8.19a   S&P futures, profit factors for 5 trend systems, 20 years through 2011.
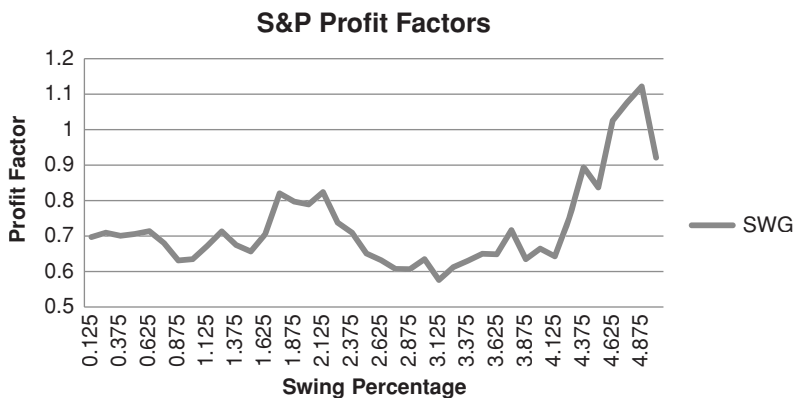
FIGURE 8.19b   S&P futures, profit factors for swing method, 20 years through 2011.

Viewing the Big Picture

When trying to decide how fast or slow to trade, it is important to recognize that shorter calculation periods result in more trades, smaller profits per trade, and more sensitivity to transaction costs, slippage, and commission. Even for the euro currency, which is a moderately trending market, fast trading lowers the profits per trade to a marginal amount, which tends to lower the profit factor as well. In Figure 8.21 the five systems are all shown in terms of the number of trades (frequency of trading) and the resulting profit factors. The chart shows that calculation periods that produced more than 200 trades in 20 years were not as good as slower choices.

Expectations

Expectations help us recognize when test results are wrong and, even more important, when actual trading varies too far from test results. Common sense goes a long way toward

Euro—Number of Trades
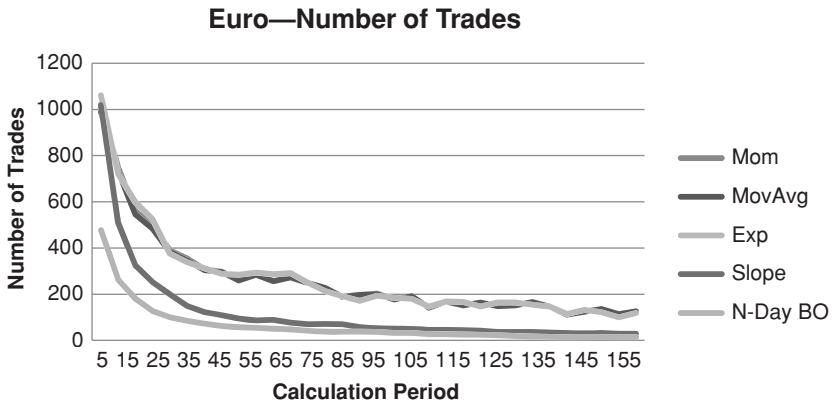
FIGURE 8.20a Euro currency futures, 20 years through 2011, number of trades.

Euro—Percentage of Profitable Trades
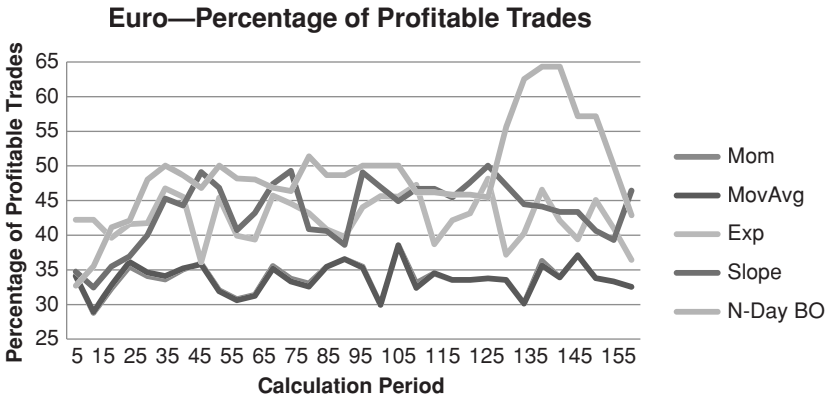
FIGURE 8.20b Percentage of profitable trades.

Euro—Average Profit per Trade
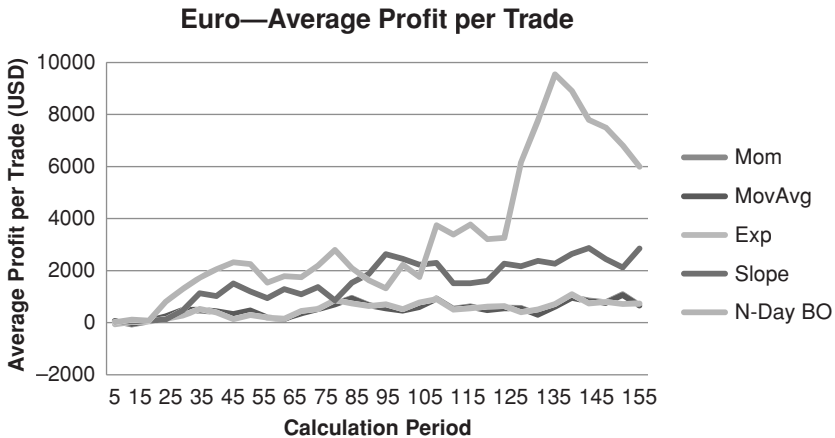
FIGURE 8.20c Average profit per trade.
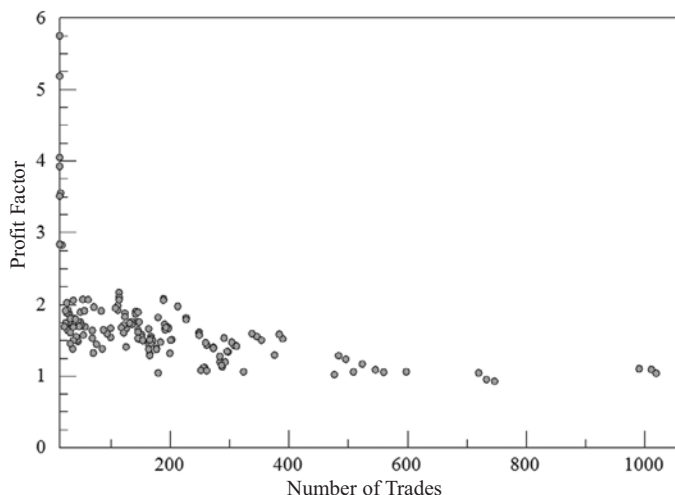
FIGURE 8.21 Euro profit factors versus number of trades, for the five strategies combined, shows that fewer trades yield much higher profit factors. Profit factors below 1.0 are losses.

keeping your work correct. Consider an engineer who is building a bridge over a large river. The data entry clerk incorrectly enters a value with the decimal shifted to the left one place. The final plans show a 10-foot bridge over a 100-foot river. Fortunately, the engineer had estimated the results and expected an answer between 80 and 120. When different systems are tested for different calculation periods, we should have a good idea of the expected results. Fast and slow trading each has its own patterns; breakout and moving averages have distinct risk profiles. Understanding these differences is important to success.

Robust Testing

To get the most value from these tests, there needs to be as much consistency as possible between tests. The following standards were used:

1. *Test periods*. All markets and all tests used the same 20-year period, ending February 2011. More data are always better, and data that include bull markets, bear markets, and sideways markets are necessary to get the best idea of how a system performs.
2. *Tests start on the same date regardless of the calculation period*. All tests should begin on the same date as the test with the longest wind-up period, in this case 160 days.
3. *Range of trend speeds*. The calculation periods should produce tests that range from only a few trades to many. It should not be necessary to fall below two or three trades per year, or more than 100 per year. Those combinations are unlikely to be traded.
4. *Use realistic costs*. Costs matter because you cannot trade for free. If your costs are too high, then you might turn a profitable system into a loss, and if the costs are too

low, then you will have unrealistically optimistic results. Try to use costs that represent actual trading.

5. *Type of results*. In the previous sections, the results were shown as net profits, profit factor, number of trades, average trade profit, and the percentage of profitable trades. If you could only choose one, then the profit factor, or information ratio, is the most important. Those statistics give you the returns relative to the risk and allow you to compare trading strategies on an equal basis.

The complete testing process is discussed in Chapter 21.

Which System Is Best?

We come back to the question that started this study of various trend-following strategies, "Which system is best?" Of the sample of markets and methods, the moving average and the *n*-day breakout seem to be the best choices but with two completely different profiles. The moving average has many more trades and many losses, with a few large profits. It is considered a conservation-of-capital approach. The breakout system tends to be more reliable and has much larger profits per trade but does it by taking very large risk. The linear regression slope seems to be somewhere in between.

Many traders prefer the returns of the breakout system but find the risk too high and the time holding the trade too long. You might think of trading a 100-day breakout as *investing* rather than *trading*. Although a moving average takes losses sooner, it holds the biggest winners for even longer. Traders are always looking for a way to take advantage of the long-term trend but not hold the trade as long and not take as much risk. This will be discussed in the next section.

The answer is "There is no best system, only trade-offs." Every investor has a different risk tolerance. Some traders need to trade a system that has a high percentage of profitable trades, and others want only small losses. It is important to start building a strategy using an underlying method that satisfies your goals.

Programs for the Six Systems

In addition to the spreadsheet that gives comparative performance of the six systems, there are individual programs on the Companion Website that will provide much more detail.

TSM Momentum

TSM Moving Average

TSM Exponential

TSM N-Day Breakout

TSM LinReg Slope

TSM Swing

The moving average program gets signals from the trendline rather than a price penetration. Each program has useful options.

TECHNIQUES USING TWO TRENDLINES

There are many situations where two trends of different calculation periods can solve a problem better than one. It is often the case that there is a dominant, long-term trend driven by government interest rate policy. Trends based on fiscal policy can last for years and can be very successful. Most traders, however, are not likely to hold a single long-term trade for the full period of its move. Even if convinced of the ultimate outcome of the trade, there can be very large price swings along the way. Most traders would rather enter and exit the market many times, in the direction of the longer-term trend, each time taking a small profit but with much smaller risk. The final result may be lower total profits, but a much more comfortable risk level for each trade.

This problem can be solved with two simple moving averages or a combination of any two trendlines of different speeds. The slower trendline, using a longer calculation period, identifies the primary trend. The faster trendline is used for timing. The faster signal does not have to be a trend at all; it can be pattern recognition or any timing method. In this section, we will use the same trending techniques previously discussed to create a system. The longer calculation period will represent the major trend and the shorter period will be used for timing. Consider the idea that a good entry point is when there is a recent short-term surge of prices in the direction of the major trend. To implement this plan, select two moving averages, one noticeably faster than the other, and apply one of the following sets of rules (also shown in Figure 8.22).

1. *Buy* when the faster moving average crosses the slower moving average going up. *Sell short* when the faster moving average crosses the slower moving average going down.

2. *Buy* when the current price crosses above *both* moving averages and close out long positions when prices cross below *either* moving average. *Sell short* when the current price crosses below *both* moving averages, and close out short positions when prices cross above *either* moving average.

3. *Buy* when the faster trendline turns up and the slower trendline is up. *Sell short* when the faster trendline turns down and the slower trendline is down. *Exit* the trade when the two trendlines are moving in opposite directions.

The first set of rules always has a position in the market, going from long to short and back again as the faster trend crosses the long-term trend. The second and third sets of rules create a neutral zone, where no position is held. Rule 2 attempts to extract the stronger part of the price move based on price, while Rule 3 looks for both trends to provide confirmation. Exiting a trade, rather than reversing, adds liquidity by reducing the order size and allows you to enter the next trade in the same direction as the previous one, instead of always reversing.

To further reduce the problem of whipsaws caused by erratic penetration of the trendlines in Rule 2, yet maintain a faster response to price change than Rule 3, a small band
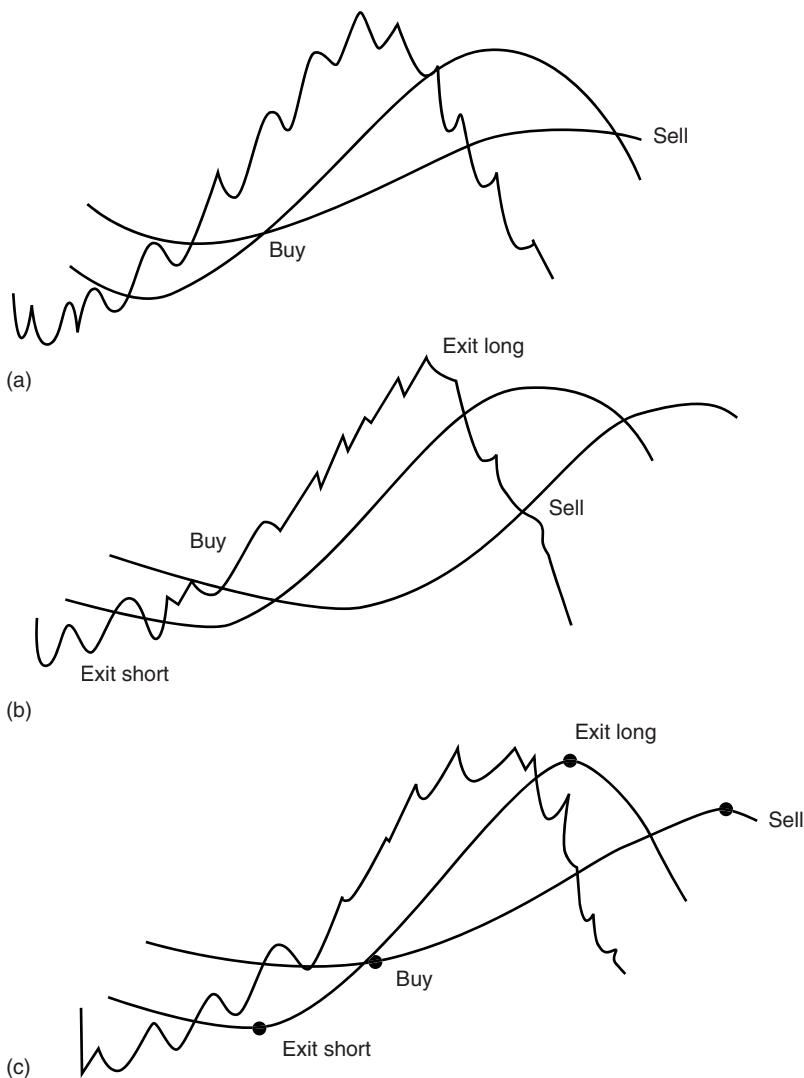
FIGURE 8.22 Three ways to trade systems using two moving averages. (a) Enter and exit when the trendlines cross. (b) Buy and sell when the price crosses the trendlines, staying out of the market when prices are between the trendlines. (c) Enter when both trendlines are moving in the same direction; exit when they conflict.

can be placed around each of the trendlines. Prices must move higher through the upper band before a buy signal occurs and then back through the lower band before that signal is reversed. It is a small safety zone that can eliminate the frequency of bad trades in proportion to the size of the band. With this technique you would want the band to be small; otherwise, you will interfere with the natural process that is the benefit of the two trendlines.

Donchian's 5- and 20-Day Moving Average System

The method claiming one of the longest recorded trading histories, beginning January 1, 1961, is *Donchian's 5- and 20-Day Moving Average.*[8] In 1961, when moving averages were considered state-of-the-art, there was less noise, and agricultural markets were the most liquid. The equivalent of a 1- and 4-week moving average would have worked well. Even now, the use of calendar periods—such as 21 and 63 days for a month and a quarter, respectively—may pick up trends driven by the action of major fund managers as they rebalance their portfolio each month, and also respond to price direction resulting from quarterly earnings reports.

Donchian's idea was to use a volatility-penetration criterion relative to the 20-day moving average, but with some added complication. The current price penetration must not only cross the 20-day moving average but also exceed any previous 1-day penetration of a closing price by at least one volatility measure. In this way Donchian places a flexible band around the 20-day trendline. One volatility measure can be calculated as the average true range over one or more days.

The 5-day moving average serves as a liquidation criterion (along with others) and is also modified by prior penetration and volatility. These features tend to make Donchian's approach an early rendition of self-adjusting rules. To maintain a human element, Donchian requires execution of certain orders to be delayed a day if the signals occurred on specific weekdays or before a holiday. The combination of different factors was the result of refinement over years of actual operation.

Rather than try to implement Donchian's idea exactly, the program *TSM Donchian Moving Average System*, available on the Companion Website, uses the calculations:

1. A 5-day moving average
2. A 20-day moving average
3. The average true range based on the longer moving average

These three calculations are then used with the rules

- If position is not long and $Close_t > MA5_{t-1} + 1ATR_{t-1}$ and $close_t > MA20_{t-1} + 1ATR_{t-1}$ then *buy*
- If position is not short and $Close_t < MA5_{t-1} - 1ATR_{t-1}$ and $close_t < MA20_{t-1} - 1ATR_{t-1}$ then *sell short*
- If position is long and ($Close_t < MA5_{t-1} - 1ATR_{t-1}$ or $close_t < MA20_{t-1} - 1ATR_{t-1}$) then exit long position
- If position is short and ($Close_t > MA5_{t-1} + 1ATR_{t-1}$ or $close_t > MA20_{t-1} + 1ATR_{t-1}$) then cover short position

---

[8]Richard D. Donchian, "Donchian's 5- and 20-Day Moving Averages," *Commodities Magazine* (December 1974).
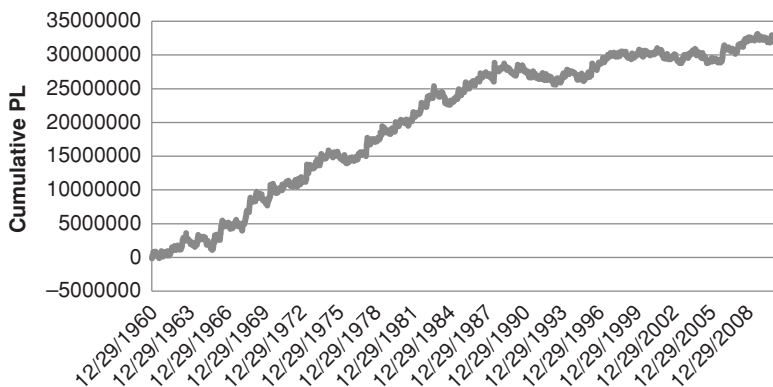
FIGURE 8.23 Donchian's 5- and 20-Day Moving Average System (somewhat modernized) applied to corn futures from 1960.

Because the price level and volatility of the market have changed dramatically since 1960, new positions should be sized according their volatility

$$Position\ size = Investment/(ATR \times Big\ Point\ Value)$$

Where ATR is calculated over the longer moving average period and the *Big Point Value* is the conversion factor for a futures contract, for example $50 for corn and $1000 for U.S. bonds.

How did this strategy perform? Applying these rules to corn, which would have been the primary market during the 1960s, and without costs (which were much higher until the mid-1990s), the cumulative profits are shown in Figure 8.23. Although the rate of return has slowed, it seems remarkable that a simple method could have been consistently profitable for 50 years. For those analysts who are interested, the program on the Companion Website allows the calculation periods to change as well as the penetration factor. Only corn was run for this example, and no parameters were tested or changed.

Donchian's 20- and 40-Day Breakout

One level slower than the 5- and 20-day average is Donchian's 20- and 40-Day Breakout. Instead of 1 week and 1 month, this looks at 1 month and 2 months. The method is far less complicated and only considers simple breakouts rather than using volatility bands. The rules are

Buy when *today's high > high* of the past 40 days

Sell short when *today's low < low* of the past 40 days

Exit longs when *today's low < low* of the past 20 days

Exit shorts when *today's high > high* of the past 20 days

Readers will recognize that this is the basis for the Turtle's trading method.

The Golden Cross and the Death Cross

The most popular stock market trending methods are the simplest, which does not mean they don't work. Of course, the 200-day moving average is shown as the key technical indicator on most financial networks, but 50, 100, and 200 days are equally popular. It is not clear how these began, but doubling the period is a simple way of keeping percentage changes the same and getting a good distribution of results over time.

The *Golden Cross* is the point at which the 50-day average crosses above the 200-day average indicating the beginning of a bullish move in the market. It has yielded very good results for the past 60 years and avoided the damaging declines of 2008. When the 50-day average crosses below the 200 day, it is ominously called the *Death Cross*.

In Figure 8.24, the results of Golden Cross are compared to the passive returns of the S&P index (SPX) and continuous futures, remembering that SPX cannot be traded. When the trend signal indicates a short sale, the 1-day returns of the 3-month T-bill rate are used for the daily returns. A spreadsheet named *TSM Golden Cross* can be found on the Companion Website.

For the 11 years beginning mid-1999, the passive return of the stock market was a loss of 7.8%. During the same period, the Golden Cross returned 66.7% using SPX and 36.7% using futures. While the cash index can be traded as the ETF SPY, the futures contract is reasonable alternative. In addition, futures can be leveraged considerably, increasing the returns (and the risk).

All the calculations are shown in the Golden Cross spreadsheet, but the way in which the returns of continuous futures are matched to SPX needs some explanation. While the daily returns of SPX are calculated as $\ln(p_t/p_{t-1})$, continuous futures are back adjusted and values can become negative. The annualized volatility of SPX was calculated in the classic way (see Chapter 2), but the annualized volatility of the futures was calculated based on the daily change in dollar value of the futures contract.
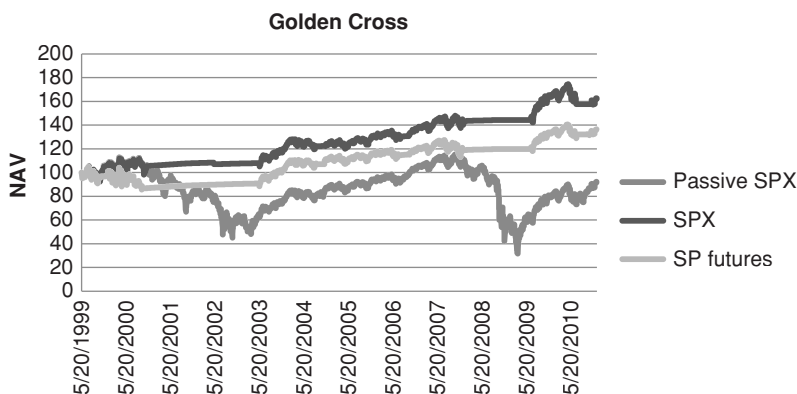
FIGURE 8.24   The Golden Cross applied to the S&P index (SPX) and continuous futures compared to the passive returns of SPX.

1. Begin with an arbitrary, but large, investment size
2. Calculate the daily returns based on the initial investment size
3. Calculate the rolling 20-day annualized volatility of the returns
4. Find the factor needed to make the annualized volatility 12%
5. Multiply the current returns by the previous volatility factor
6. Create the NAVs from the volatility-adjusted returns

In Figure 8.24 the annualized volatility of SPX is 11.5%, and the volatility of the futures returns is 10.9%. The process of adjusting the portfolio volatility is discussed in detail in Chapters 23 and 24.

ROC Method

Another classic method for trading the major index is *Woodshedder's long-term indicator*.[9] The rules are

- *Buy* when the 5-day ROC (rate-of-change) is below the 252-day ROC for two consecutive days.
- *Exit the long* when the 5-day ROC is above the 252-day ROC for two consecutive days.
- When there is no position, the system earns one-half of the cash 3-month T-bill rate.

In Figure 8.25, the results are compared to the Golden Cross, both using SPY as the basis for trading, from April 15, 1994 (the first available data for SPY plus the 252-day windup), through September 2011. No costs were charged although SPY has administrative cost included. The ROC method far outperformed the Golden Cross even with an annualized volatility of 14.1% compared to the Golden Cross volatility of 10.5%. Note that trading the SPY gave better returns than SP futures, shown in Figure 8.25.

Staying Ahead of the Crowd

There is always an attempt to find out where others are placing their orders and get ahead of them. For example, if you know that most trend-followers are using a 30-day calculation period, then using a 28-day average might edge them out. During the 1980s and 1990s there was a trend system that used 8 and 18 days to beat the 10 and 20 days that was most popular. The following calculations would use *fastperiod* = 8 and *slowperiod* = 18. The *differenceperiod* = 9.

$$FasterAverage = Average(close, fastperiod)$$
$$SlowerAverage = Average(close, slowperiod)$$

---

[9]The *Woodshedder* blog can be found at www.ibankcoin.com/woodshedderblog and covers many other strategies. This method was reviewed by *MarketSci blog* on October 4, 2011, but used SPX (the cash index) rather than SPY.
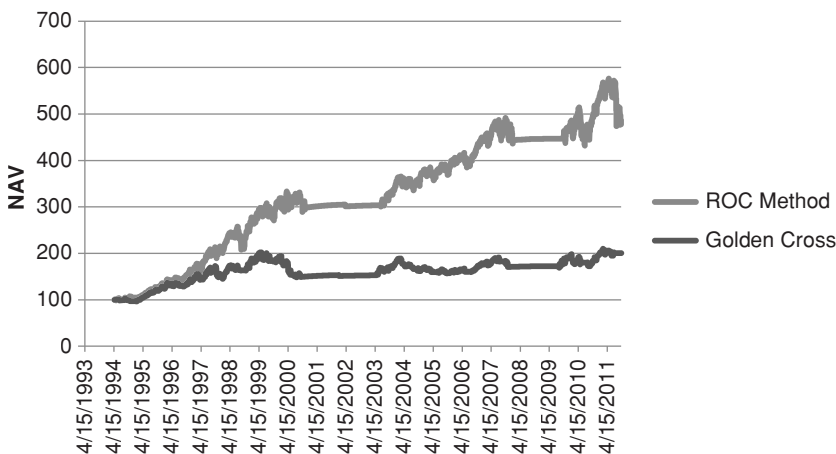
FIGURE 8.25   Comparison of Golden Cross and ROC Method using SPY.

$$TrendDifference = FasterAverage - SlowerAverage$$
$$DifferenceAverage = Average(TrendDifference, differenceperiod)$$

The trading rules were

Buy when *today's TrendDifference > yesterday's DifferenceAverage*
Sell short when *today's TrendDifference < yesterday's DifferenceAverage*

Although these calculation periods may not be profitable in today's markets, the idea of being slightly ahead of the crowd can give you *free exposure*, a small jump in profits caused by many orders following yours. If you can figure out where the crowd is buying and selling, then this concept will give you an edge.

MULTIPLE TRENDS AND COMMON SENSE

If two trendlines can improve trading, it should follow that three or more are even better—but there may be more problems than benefits. Many analysts subscribe to the idea that *simpler is better*. A single moving average may not have a high percentage of profitable trades but the longer-term periods work because they capture the fat tail.

With the use of two trends, the number of combinations expands rapidly. Is there a *best* relationship between the slower speed and the faster one, that is, should the faster trend period be ¼ of the slower? Certainly, a 38- and 40-day combination will not offer much value, but is a 10-day and 40-day the right combination?

Consider something else. If a 10-day trend is not profitable, and a 20-day trend is not profitable, each taken on their own, but the combination is profitable, would you trade

TABLE 8.7 Comparison of 1 and 2 Trend Systems, Eurodollar Interest Rate Futures, 1990–2011

Trend	Number of Trades	Days Held	Total Profit	Ratio	Avg Trade Profit	%Prof Trades
60	164	31	16537	1.65	149	22.5
20	350	16	21737	1.44	62	31.0
Both	254	16	15310	1.50	92	34.6

it? These questions will be addressed in Chapter 21, but some of the concepts should be clear now. If, by computer testing, we were to "discover" that a combination of 2 or 3 trends was profitable, would you be convinced to trade it? Not likely, unless a very large percentage of the combinations were profitable or each trend served a particular purpose and had a calculation period that reflected that purpose.

For example, if longer-term trends are intended to track macro fundamental policy and were generally profitable, then that longer-term period would be a good candidate for one of the trends. But those trades can be held for months, and you don't like that profile, so you want a shorter trend to tell you when to get in and out of that trade, always holding a position in the direction of the long-term trend. Using Eurodollars with 60-day and 20-day trends over the past 20 years, with a $40 round-turn cost, Table 8.7 shows the change in performance. Both single trends do very well, but the 60-day holds trades for an average of 31 days, about 1½ months. The 20-day trend would be better, but the average profit per trade is only $62. By combining the two trends and trading only in the direction of the long-term trend, the profits per trade jump to $92, and the days held remain the same as the 20-day trend. Overall, you get slightly worse performance than the long-term trend but hold the trade for half the time. That makes sense because you are only trading the longs rather than both longs and shorts.

Another benefit of the two-trend method is that you are in the market only 50% of the time. That reduces your risk, especially the risk of a price shock. That is a benefit not to be taken lightly.

Three Trends

Is there a rationale for more than two trends? If the long-term trend is for market direction, and the shorter one is to reduce the length of the holding period, then the third could be using for entry timing. The third trend could be very fast, perhaps 3 days.

Gerald Appel[10] adds three rate-of-change (ROC) indicators together (actually momentum, the difference between the price today and the price n days ago), and applies the composite to the S&P index (SPX), all expressed in percent. He recommends buying

---

[10] Gerald Appel, *Technical Analysis: Power Tool for Active Traders* (Upper Saddle River, NJ: FT Prentice Hall, 2005), 59.

when the composite crosses above 4% and exiting when it falls below 4%. If you consider the upwards bias in the S&P, the 4% threshold may not be arbitrary.

Modified 3-Crossover Model

The justification for using three trends is that one or more slower moving averages may result in a buy or sell signal at a time when the prices are actually moving opposite to the position that is about to be entered. This may happen if the trading signal is generated when the two moving average lines cross, rather than when price move through the averages. The slope of a third, faster moving average, $MA3_t - MA3_{t-1}$, can be used as a confirmation of direction to avoid entry into a trade that is going the wrong way. This filter can be added to any moving average or multiple moving average system with the following rule:

> *Do not enter a new long (or short) position unless the slope of the confirming moving average (the change in the moving average value from the prior day to today) was up (or down).*

The speed of this third, confirming moving average only makes sense if it is equal to or faster than the fastest of the trends used in the Crossover System. A program to test the three-trend model is *TSM Modified 3MA Cross*, available on the Companion Website.

In earlier tests of the 3-Crossover method compared to the 2-Crossover, results showed that the added timing in the 3-Crossover reduced the number of trades and increased the size of the returns per contract. Overall, the profitability remained about the same.

4-9-18 Crossover Model

During the late 1970s, the 4-9-18 Crossover model was very popular. It seems likely that the selection of 4, 9, and 18 days was a conscious effort to be slightly ahead of the 5, 10, and 20 days frequently used in moving average systems during that period. It is also likely to have been the outcome of the first computerized testing. Even now, high frequency traders continue to look for the smallest edge that keeps them ahead of the competition. In addition to the marginally faster calculation periods each moving average is (nearly) twice the speed of the prior, enhancing their uniqueness for recognizing different trends. Increasing the period in this way keeps a constant percentage difference.

To get an idea of how three trends compare to either one or two trends, a small test was run using Eurobund futures from 1990 to mid-2011. Granted this is a trending market, and one test does not reflect the big picture; nevertheless, the results, shown in Table 8.8, were unexpected. Because we already know that faster trends are not performing as well in recent years as they did in the 1970s, the test of Eurobunds compared a 40-day and an 80-day calculation period with a 40-80-day crossover and a 20-40-80-day combination. The 3-crossover method outperformed all others but also reduced the number of trades by emphasizing the trend direction.

TABLE 8.8 Comparison of Single Trend Results with a 2- and 3-Moving Average Crossover Strategies, Applied to Eurobunds, 1990 to Mid-2011

Strategy	Net PL	Profit Factor	Trades	Profit/ Contract
40-Day MA	13300	1.11	310	42
80-Day MA	55620	1.88	152	365
40-80 Crossover	35040	1.45	229	153
20-40-80 Crossover	61210	2.80	42	1457

COMPREHENSIVE STUDIES

Because computerized testing platforms have made it easy for anyone to test any number of trends in combination, there have been very few comprehensive studies published since 1990. The exceptions are Colby's *The Encyclopedia of Technical Market Indicators* and Bulkowski's *Encyclopedia of Chart Patterns*, both of which show results in a standard form and make it easy to compare the differences between systems. But most traders seeking a strategy will need to test it themselves and add their own special features. Both Colby and Bulkowski will give you a good idea of which methods and patterns to avoid.

There is a great deal to learn from putting the results of various systems and markets side by side. Earlier in this chapter, there is an informative comparison of six major trending systems; in Chapter 21 tests of single-trend strategies are compared for a standard set of 17 futures markets. In addition, a 2-trend crossover strategy is compared to the single trend methods. An objective of the testing process is to find parameters that succeed over time; this discussion can also be found in Chapter 21.

SELECTING THE RIGHT TREND METHOD AND SPEED

Up to now, the selection of the right moving average, the one that will work in the future, has only been discussed in general terms. The success of a single calculation period for a single trend strategy does not mean that it is the right choice for trading. In addition, the best moving average speed for institutional or commercial participants may be very different from that of an active trader.[11] For example, a mutual fund receives new investments that must be moved into the market, collectively, once or twice each month. In the same way, a cattle feed lot will choose one time each month to fix the price of new inventory. A 3-day moving average might generate 5 to 10 buy and sell signals in one month, each the result of a 2-day price move—an ineffective tool for either participant

---

[11] Perry J. Kaufman, "Moving Averages and Trends" in Todd Lofton, ed., *Trading Tactics: A Livestock Anthology* (Chicago: Chicago Mercantile Exchange, 1986).

looking for just one place to enter the market. A calculation period of 10 days may come closer to generating the one buy signal needed for the fund or one sell signal that is best for the hedger. The noncommercial trader is not concerned about the frequency of trades, only the returns and risk.

For the trader or speculator, the right moving average speed is the one that produces the best performance profile. This profile could be simply maximum profits, or it could be a more complex combination of profits, risk, and time in the market. In Chapter 21, automated testing is used to find the combination of speed, stop-loss, and other rules that best satisfy an objective; a computer, however, is not always necessary.

Computer testing of a trend system or other trading strategies sometimes leads to solutions that are highly fitted. The computer may find that a 3-day moving average was slightly more profitable and had lower risk than a 20-day average. Our common sense tells us that the results of the 3-day system will be more difficult to attain in real trading because execution costs will have a larger impact. An occasional fast market may cause the execution price to be far off from the price indicated by the system signal. A slower trend selection with fewer trades is less affected by poor executions.

Dominant seasonal factors are an important influence on the calculation period of the trend. While some stocks, such as travel and leisure, can be highly seasonal, their seasonal price patterns can be overwhelmed by a strong trend in the overall market, as measured by the S&P 500. In fact, the arbitraging of the S&P 500 futures with the actual stocks has significantly changed the patterns of the many stocks, forcing many of them to have higher correlations. However, a grain trader knows that the price pattern has a clear cycle each year. At best, we can expect one long upwards move followed by a shorter, faster decline. Not all trend speeds can capture the profits in this pattern. For example, if the uptrend lasts for 6 months, a 6-month moving average will not see any of it; therefore, it is necessary to use a moving average with a period less than one quarter of the length of the trend. If computerized testing of a large range of moving average calculation periods results in a "best" moving average period of six months or more, that choice should be interpreted as a failure to capture the seasonal move.

We also discussed the "right" trend method earlier in this chapter. Tests over many years and many diverse markets will show that the differences in net returns using different trend strategies are small. Those differences will be larger when the trend speeds are faster. The most important differences are not in the profits but in the frequency of trades, the size of the individual returns, and the risk of each trade. Experience shows that the primary reason why trend-following systems work is because sustained price moves exist, driven mostly by government interest rate policy. Every trend-following system can capture these moves.

Selecting the Trend Speed

Chapter 21 will discuss the systematic ways to find the "best" trend, and that is the most likely way that traders make their choice. But you can find a reasonable choice by simply looking at a price chart. The trends that two traders see are often different. Some traders

FIGURE 8.26 Finding the moving average calculation period from the average span of peaks and valleys.

immediately focus on long-term price trends; others see much shorter movements. To find your own best trend speed without the use of a computer, mark on a chart the beginning and end of each price move that you would like to capture. These trends may occur every few days, or only three or four times each year.

Using a daily price chart of IBM from July 2000 through May 2002 (Figure 8.26), the tops and bottoms of the major price swings that we would like to capture were circled. Noting that each gridline on the chart represents one month, there were eight tops and bottoms over a total period of 22 months. The average price swing was then 2¾ months, nearly one calendar quarter. Because it is so close to a quarterly value, which is the period of earnings reports, we will choose three months as the average swing period. Applying the rule of thumb that the trend can be isolated using a calculation period of one-quarter of the swing period, the moving average period becomes 16 days.

Applying a 16-day moving average system to the two years of IBM price data, buying and selling whenever the trendline changes direction, the trade results are shown in Table 8.9. There are a total of 36 trades, of which 20 were profitable (no transaction costs were used)—a very high percentage for a trend-following system. Total profits were better than $114 per share, on an average share price of about $100, showing that it is not necessary to optimize using a computer to create a successful trading program. Of course, there is no assurance that this pattern of swings will continue, but basing the decision on quarterly swings, which corresponds to earnings reports, is a hopeful sign.

TABLE 8.9 Performance of 16-Day Moving Average Applied to IBM from July 2000 through May 2002

Trade #	Entry Date	Entry Order	Entry Price	Exit Date	Exit Order	Exit Price	Profit/ Loss	Total P/L
1	6/29/2000	Sell	114.00	7/20/2000	SExit	117.25	−3.25	−3.25
2	7/20/2000	Buy	117.25	7/24/2000	LExit	112.50	−4.75	−8.00
3	7/24/2000	Sell	112.50	7/25/2000	SExit	112.00	0.50	−7.50
4	7/25/2000	Buy	112.00	9/18/2000	LExit	123.25	11.25	3.75
5	9/18/2000	Sell	123.25	11/9/2000	SExit	99.44	23.81	27.56
6	11/9/2000	Buy	99.44	11/10/2000	LExit	93.00	−6.44	21.12
7	11/10/2000	Sell	93.00	11/13/2000	SExit	97.44	−4.44	16.68
8	11/13/2000	Buy	97.44	11/27/2000	LExit	98.44	1.00	17.68
9	11/27/2000	Sell	98.44	12/5/2000	SExit	103.38	−4.94	12.74
10	12/5/2000	Buy	103.38	12/6/2000	LExit	96.75	−6.63	6.11
11	12/6/2000	Sell	96.75	1/5/2001	SExit	94.00	2.75	8.86
12	1/5/2001	Buy	94.00	2/20/2001	LExit	111.50	17.50	26.36
13	2/20/2001	Sell	111.50	4/5/2001	SExit	98.21	13.29	39.65
14	4/5/2001	Buy	98.21	5/11/2001	LExit	111.81	13.60	53.25
15	5/11/2001	Sell	111.81	5/15/2001	SExit	113.58	−1.77	51.48
16	5/15/2001	Buy	113.58	5/23/2001	LExit	117.40	3.82	55.30
17	5/23/2001	Sell	117.40	5/24/2001	SExit	119.60	−2.20	53.10
18	5/24/2001	Buy	119.60	5/29/2001	LExit	115.27	−4.33	48.77
19	5/29/2001	Sell	115.27	6/5/2001	SExit	116.97	−1.70	47.07
20	6/5/2001	Buy	116.97	6/12/2001	LExit	117.25	0.28	47.35
21	6/12/2001	Sell	117.25	6/22/2001	SExit	112.87	4.38	51.73
22	6/22/2001	Buy	112.87	6/25/2001	LExit	112.65	−0.22	51.51
23	6/25/2001	Sell	112.65	7/31/2001	SExit	105.21	7.44	58.95
24	7/31/2001	Buy	105.21	8/6/2001	LExit	106.51	1.30	60.25
25	8/6/2001	Sell	106.51	8/10/2001	SExit	104.95	1.56	61.81
26	8/10/2001	Buy	104.95	8/17/2001	LExit	104.59	−0.36	61.45
27	8/17/2001	Sell	104.59	10/5/2001	SExit	98.02	6.57	68.02
28	10/5/2001	Buy	98.02	1/11/2002	LExit	120.31	22.29	90.31
29	1/11/2002	Sell	120.31	3/4/2002	SExit	105.90	14.41	104.72
30	3/4/2002	Buy	105.90	3/6/2002	LExit	106.30	0.40	105.12
31	3/6/2002	Sell	106.30	3/12/2002	SExit	108.50	−2.20	102.92
32	3/12/2002	Buy	108.50	3/26/2002	LExit	102.90	−5.60	97.32
33	3/26/2002	Sell	102.90	5/22/2002	SExit	84.00	18.90	116.22
34	5/22/2002	Buy	84.00	5/24/2002	LExit	83.10	−0.90	115.32
35	5/24/2002	Sell	83.10	5/28/2002	SExit	82.08	1.02	116.34
36	5/28/2002	Buy	82.08	5/31/2002	LExit	80.45	−1.63	114.71

Another approach to finding the trend period is to consider the worst price retracement. In the IBM chart, the move from $87 to $105, beginning in November 2000 and lasting about two months, is the one to avoid. Remembering that a trend is neutralized with regard to a price move when the trend period is the same length as the total move, we apply a moving average of 42 days. This method successfully avoids the price correction

and holds the downtrend, but the much slower trend nets a significant loss over the 2-year period. It is best to accept the frequent small losses that are natural in a trend-following system than attempt to remove them.

MOVING AVERAGE SEQUENCES: SIGNAL PROGRESSION

Consider the case where you have selected a 20-day moving average to trade. You enter the day long Biotech, and you get a sell signal. However, you are unaware that the 19-day and 21-day moving averages did not get sell signals. This means that the day that was dropped off the calculation 20 days ago caused a slight shift not seen by the other neighboring trends. This can be an important piece of information when assessing the reliability of the trend signal.

A moving average is simply a consensus of direction. It is an approximation of values intended to steer a trader to the right side of the market at the right time. It is most fallible when prices are changing direction or going sideways. Any information that clears up the problem is helpful. For any trend system, it is best to see a steady progression of trend changes from the short term to the long term. This is seen in the following tables, where u is an uptrend and d is a downtrend associated with the calculation period above those letters.

In Table 8.10, prices have turned up in such a way that the trend calculation periods 1 through 19 show uptrends, while calculation periods from 20 and higher have not yet turned. Unfortunately, normal price movement is not often as uniform as this example. The shorter-term trends can be very erratic, and often appear in smaller, alternating groups of up and down trends (see Table 8.11). This is easily explained because adding and subtracting one day when only two, three, or four days are used in the moving average calculation can quickly change the direction of the trend. As you get to longer intervals, such as 20, 30, and 50 days, this is not the case, and in reality, it does not happen often. Yet when it does, the trend change is not to be trusted.

There are also cases where the longer trend begins to reassert itself and the results appear the same as in Table 8.10; however, the trend change occurs from the longer-term down (from right to left instead of left to right). The case we must watch for satisfies neither of these, but occurs in an erratic pattern, such as in Table 8.12. Here we see a dominant long-term uptrend with the very short end turning down. Because of another downturn a few days ago, which then disappeared, this most recent downturn also caused a

TABLE 8.10 Orderly Trend Change

Moving Average Period in Days																								
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Trend u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	u	d	d	d	d	d	d

TABLE 8.11 Erratic Trend Change for the Short Calculation Periods

	Moving Average Period in Days																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Trend	u	u	d	d	u	d	u	u	u	u	u	u	u	u	u	u	u	u	u	d	d	d	d	d	d

shadow turn in the 20-day range. Is it a leading indicator or a false signal? All indications are that smooth changes in a trend are more reliable precursors of change. Another case is given in the second line of Table 8.12. Here, the smooth trend change from up to down is occurring from left to right; however, as it gets to 13 days, it also jumps ahead to 19 and 20 days, leaving days 14 through 18 still in an uptrend. For trends in this faster range, it appears best to wait for all fastest trends to change. As the calculation period becomes longer, it is unrealistic to expect all faster trends to be the same; therefore, you will need to settle for an orderly change in a group of trends faster than the target trend period.

An example of this process is shown in Figure 8.27. Moving average calculation periods of 5 to 50 days are shown in increments of five days for a total of 44 consecutive days. This illustration points out how the long-term uptrend (X) is breached by shorter-term, less consistent trends. Perhaps the best trend is the one with the majority of *X*s or *O*s on the same line.

Averaging the Sequences

The idea of requiring consistency in a range of trends can be automated by selecting a range of calculation periods preceding a target period, finding the trend signal (an uptrend or downtrend), and then deciding according to one of two rules:

1. Average the final trend values to get the average trend result. Compare the previous average result with the current value to determine the direction of the trend.
2. Scan the trend directions for consistent progression.

When selecting the range of calculation periods, start from 1 if the target period is small (e.g., your intended trend is 15 days). If you are looking at an intermediate trend period, for example, 30 days, you may want to include the range from 20 to 30, or 20 to 32. A few up-and-down price moves that make the short-end erratic should not alter the medium-term trend direction. By extending the calculation periods slightly past the target period, you gain confirmation at the cost of a small lag.

TABLE 8.12 Progression of Trend Changes

	Moving Average Period in Days																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Trend	d	d	d	d	u	u	u	u	u	u	u	u	u	u	u	u	u	u	d	d	u	u	u	u	u
Trend	d	d	d	d	d	d	d	d	d	d	d	d	d	u	u	u	u	u	d	d	u	u	u	u	u

					Moving Average Period					
Day	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**
1	X	X	X	X	X	X	X	X	X	X
2	O	O	X	X	X	X	X	X	X	X
3	O	X	X	X	X	X	X	X	X	X
4	X	X	X	X	X	X	X	X	X	X
5	X	X	X	X	X	X	X	X	X	X
6	X	X	X	X	X	X	X	X	X	X
7	X	X	X	X	X	X	X	X	X	X
8	O	X	X	X	X	X	X	X	X	X
9	X	O	O	O	X	X	X	X	X	X
10	O	O	O	O	O	X	X	X	X	X
11	O	O	O	O	X	X	X	X	X	X
12	O	O	O	O	O	O	O	X	X	X
13	O	O	O	O	X	X	X	X	X	X
14	X	O	O	O	O	O	X	X	X	X
15	O	O	O	O	O	O	O	O	O	X
16	O	O	O	O	O	O	O	O	O	O
17	O	O	O	O	O	O	O	O	O	O
18	O	O	O	O	O	O	O	O	O	O
19	X	O	O	O	O	O	O	O	O	O
20	X	O	O	O	O	O	O	O	O	O
21	X	X	X	O	O	O	O	X	X	X
22	X	X	X	X	O	X	X	X	X	X
23	X	X	X	O	O	O	O	X	X	X
24	X	X	X	X	O	O	O	X	X	X
25	X	X	X	X	O	O	O	X	X	X
26	O	X	X	O	O	O	O	O	O	X
27	O	O	O	O	O	O	O	O	O	O
28	O	O	O	O	O	O	O	O	O	O
29	O	X	X	X	O	O	O	O	O	X
30	X	X	X	X	X	O	O	O	X	X
31	X	X	X	X	X	O	O	O	O	X
32	X	X	X	X	X	X	X	X	X	X
33	X	X	X	X	X	X	X	X	X	X
34	X	X	X	X	X	X	X	X	X	X
35	X	X	X	X	X	X	X	X	X	X
36	X	X	X	X	X	X	X	X	X	X
37	X	X	X	X	X	X	X	X	X	X
38	O	X	X	X	X	X	X	X	X	X
39	O	O	X	X	X	X	X	X	X	X
40	O	O	X	X	X	X	X	X	X	X
41	O	X	X	X	X	X	X	X	X	X
42	X	X	X	X	X	X	X	X	X	X
43	X	X	X	X	X	X	X	X	X	X
44	X	X	X	X	X	X	X	X	X	X

FIGURE 8.27 Sequences of moving averages.

The idea that trends are sensitive to small price changes and the drop-off effect is not a surprise. An alternative to examining sequences is simply to select a number of calculation periods and net out the trends looking for a consensus. That would remove any dependence on a single selection.

EARLY EXITS FROM A TREND

By now we know that capturing the fat tail is necessary for the success of trend-following systems. However, there are always practical exceptions if you are allowed to add discretion to your trading decisions. One of the oldest truths for trend-following is "Take your losses and let your profits run." By imposing profit-taking, or even stop-losses, this can be changed to "Take your profits and let your losses run." There is a need to be very careful when making exceptions. But consider the following situation.

Interest rates have declined for nearly all of the 30 years from 1981 through 2011. For many traders, that's more than their entire professional career. To profit from this move, a slow trend system can track the 10-year Treasury note futures contract, a municipal or corporate bond index, or any number of varying maturity funds. If a 200-day trend were used, then there would be a lag of 100 days. That is, for a bond fund, the current value of the trendline would reflect the bond prices at the midpoint of the calculation period, 50 days. If the yield on interest rates had steadily dropped a total of 2% during the past year, then the trendline would be lagging about 1% behind current yields. That can translate into a large loss in unrealized gains.

One advantage of macrotrends is that they are based on a sustained economic policy. If that policy changes, then the trend is over, even if the trendline has not yet reversed direction. If the Fed were now to raise interest rates (or strongly hint that a rate hike is likely), it would signal a shift in policy. You can reasonably conclude that basis for the long-term uptrend in prices is over and that the trendline is due to turn down. A central bank rarely raises rates one month and lowers them the next. Because the very slow trend lags far behind the actual market price, it may be six months before the trendline actually signals a change of position. This will occur after a large part of your profits has been given back to the market. Exiting the trade when the fundamentals change would be a safe way of capturing more profits and being exposed to less market risk. Often, these decisions are clear only after the fact. In 2010 it seemed that Fed policy was going to change, yet 2011 was one of the strongest trend years in history, with yields reaching record lows.

MOVING AVERAGE PROJECTED CROSSOVERS

If moving averages can successfully be used to identify the trend direction, it follows that a projection of the moving average will be valuable in anticipating when the trends

will change. If a moving average trading strategy uses a single trend, the forecasted price (*CP*1) at which the standard *n*-day moving average line would cross is

$$CP1 = \frac{\text{Sum of last } N-1 \text{ prices}}{N-1} = \frac{\displaystyle\sum_{t-N+2}^{t} P_i}{(N-1)}$$

That is, an *n*-day moving average would cross the next price at the value equal to the *n*–1-day moving average.

The price (*CP*2) needed to cause two moving averages, of periods *m* and *n*, to cross is[12]

$$CP2 = m \times \left(\frac{\sum most\ recent\ m-1\ prices}{m} - \frac{\sum most\ recent\ n-1\ prices}{n} \right)$$

The projected crossover price is most useful when it is likely that a trend change will occur within a few days, that is, when the two moving averages begin converging and become close in value. Acting on the expected price would give the trader a great advantage in order execution. A chart of this, however, may not appear to be much different from a simple *relative strength* indicator. The difference between the price and the moving average line constitutes the relative strength.

The change in the projected crossover is considered a more valuable tool by Lambert.[13] He creates a *Market Direction Indicator* (MDI) with the following formula:

$$MDI = \frac{100 \times (Crossover\ price_{\text{previous}} - Crossover\ price_{\text{today}})}{Average\ of\ past\ 2\ day's\ prices}$$

The point at which the MDI crosses the zero line moving higher is a buy signal, and the point where it crosses moving lower is a sell signal.

Forecasting When the Moving Average Will Turn

There are two basic rules for generating a moving average signal, when the price penetrates the moving average trendline and when the moving average trendline changes up or down. In the previous section, the forecast was based on when the price crosses the moving average, which is a very common way of generating a trading signal.

Tests often show that results are better when the trendline itself is used to generate the signal rather than the penetration. This was discussed earlier in the chapter, but the rationale goes that once prices have been included in the moving average, then it is the moving average that tells you the correct direction, and that the price penetration takes

---

[12] Calculation courtesy of Alexander Solodukhin, Mizuho Alternative Investments, New York.

[13] Donald R. Lambert, "The Market Directional Indicator," *Technical Analysis of Stocks & Commodities* (November/December 1983).

away from the value of the trendline. There are two reasons for this. The trendline is smoother than price movement, so the results are more uniform, and the price penetrations generate many more trades and additional costs.

Having decided to use the trendline, then how do you forecast when that line will change direction? It is simply whether today's price is greater or less than the price that is dropped off the end of the moving average period, n days ago. And, that being the case, using the moving average trendline is exactly the same as using the n-day momentum which states that you go long when today's price is greater than the price n days ago. It seems too simple, but in most cases the results are better than using the price penetration of a moving average.

Momentum and Oscillators

The study of momentum and oscillators is the analysis of price changes rather than price levels. Among technicians, momentum establishes the speed of price movement and the rate of ascent or descent. Analysts use *momentum* interchangeably with *slope*, the angle of inclination of price movement usually found with a simple least squares regression (Figure 9.1). In mathematics it is also the *first difference*, the difference between today's price and the previous day. Momentum is often considered using terms of Newton's Law, which can be restated loosely as *once started, prices tend to remain in motion in more-or-less the same direction.*

Indicators of change, such as momentum and oscillators, are used as leading indicators of price direction. They can identify when the current trend is no longer maintaining its same level of strength; that is, they show when an upwards move is decelerating. Prices are rising, but at a slower rate. This gives traders an opportunity to begin liquidating their open trend trades before prices actually reverse. As the time period for the momentum calculation shortens, this indicator becomes more sensitive to small changes in price. It is often used in a *countertrend*, or *mean reversion* strategy. The *change in momentum*, also called *rate of change*, *acceleration*, or *second difference*, is even more sensitive and anticipates change sooner.

Before beginning a discussion of various momentum calculations, a brief comment on terminology is necessary to understand how various techniques are grouped together. The use of a single price, such as Microsoft at $25.50 or gold at $1400, has no direction or movement implied. We are simply relating a price level and not indicating that prices are going up or down.

Next, we describe the speed at which prices are rising or falling. To know the speed, it is not enough to say that the S&P rose 3 points; you must specify the time interval over which this happened—"the S&P rose 3 points in 1 hour." When you say that you drove your car at 60, you really imply that you were going 60 miles per hour, or 60 kilometers per hour. This description of speed, or distance covered over time, is the same information that is given by a single momentum value. Then, if the *daily* momentum of the NASDAQ 100 is +10, it is rising at the rate of 10 points per day.
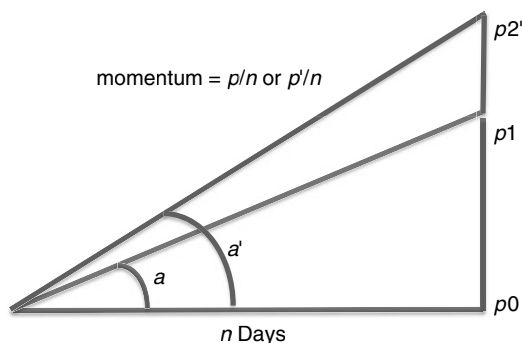
FIGURE 9.1   Geometric representation of momentum.

Having made a point of saying that momentum is change over time, which it is, the "industry" uses *momentum* to mean *price change* with the time implied (and often not even given) where *speed* is always the price change divided by the time interval. The same interpretation will be used here. The term *rate of change* (ROC) will also be used to describe *acceleration*, an increase or decrease in the speed; when the rate of change is zero, we are talking about speed and momentum.[1]

You will see in this chapter that acceleration is more sensitive than speed. We will begin with the least sensitive indicator, momentum, increase to acceleration, and then address the variations and applications, including *divergence*, the most popular use of momentum indicators.

MOMENTUM

Momentum is the difference between two prices taken over a fixed interval. It is another word for *speed*, the distance covered over time; however, everyone uses it to mean *change*. For example, today's 5-day momentum value M would be the difference between today's price p_t and the price five days ago:

$$\text{5-day momentum, } M = p_t - p_{t-5}$$

Using notation familiar to many programmers:

$$\text{5-day momentum} = \text{price} - \text{price}[5]$$

---

[1]Be careful about terminology. Many books and software use rate of change (ROC) interchangeably with momentum, the change in price over n days, rather than the change in momentum. Also, momentum indicators can also be used for mean reversion and may be called *oscillators* because their values swing from positive to negative.

or

$$n\text{-day momentum} = \text{price} - \text{price[n]}$$

where the notation [n] refers to the price n days ago.

The momentum value M_t increases as the change in price increases over the same 5-day period. Figure 9.1 shows how the momentum changes as the price increases over the same time period. Over n days, the price moves from p_0 to p_1. It forms angle a and has a momentum of $p_1 - p_0$. If prices had increased to point p_2, then the momentum would have been greater, and the angle a′ would have been larger. For example, if the 5-day change in price is an increase of 100, then the momentum is simply 100, but the speed is 100/5 = 20. If prices increase by 100 points over a 10-day time interval, the momentum would still be 100, but the speed, or slope, would be 100/10 = 10. When calculating momentum, the interval of the calculation must always be stated. To avoid confusion, we state that the *5-day momentum is 20* and the *10-day momentum is 10*.

Today's 5-day momentum, which we will now show as $M(5)_t$ can range in value from the maximum upwards move to the maximum downwards move that the price can make in five days; the momentum is zero if prices are unchanged after five days. For stocks, there is no actual limit on the maximum price range over any time interval. In cases such as Enron, prices could collapse to zero in short order. From a practical view, most stocks and futures markets have a history of volatility that relates to their price level. The higher the price, the greater the price moves. As the price of gold rose to $1000 per ounce, the 5-day momentum could have been as high as $100 per ounce ($20 per day). When it was at $250 and investors lost interest, it may have shown a $1 per ounce change over five days.

Momentum is not volatility. Gold can move from $1200 to $1250 in two days then back to $1200 over the next three days and the momentum would be zero but the volatility would be high.

Figure 9.2 shows a typical pattern of momentum. Smooth prices are used to make the relationship between price change and momentum clear. During the first 15-day cycle, prices rise steadily, peaking on day 9. As prices rise at a slower rate during days 5 through 9 the momentum declines but remains above zero. As long as the 5-day momentum is greater than zero, prices are rising. The highest momentum value occurs on day 3 when prices have the largest positive price change.

On day 10 the momentum value is zero. Looking at the price chart, this occurs when the 5-day change in price is zero. The lowest momentum value occurs on day 16 when the prices have declined the most. After prices reach their lowest point and begin up, the momentum is negative but rising.

If prices always moved in the smooth pattern shown in Figure 9.2, then momentum would be a perfect leading indicator. You could buy when momentum turned up and sell when it turned down, and be ahead of the change in price direction. Unfortunately, prices movement is irregular; consequently, momentum values are irregular. Because momentum contains valuable information, we will look at a number of systematic ways to evaluate momentum and improve trading. These focus on extreme momentum values
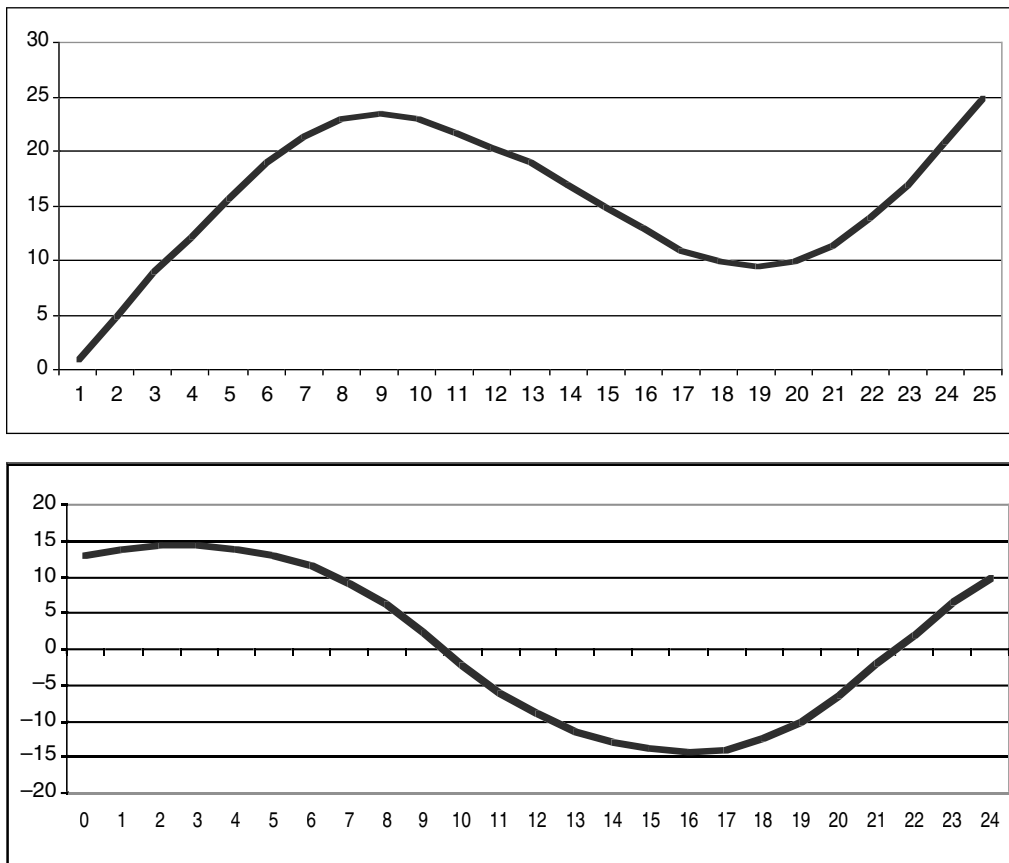
FIGURE 9.2   Price (top) and corresponding momentum (bottom).

and momentum patterns, all of which have proved to be valuable additions to trading strategies, in particular to the timing of entries and exits.

Characteristics of Momentum

Momentum is a way of smoothing price movement and can serve the same purpose as a trend. Although the momentum values are not as smooth as a moving average, larger momentum periods reduce the extremes seen in the price chart. One key advantage of momentum is that it does not have the lag that exists in a moving average.

The AOL chart (Figure 9.3) compares 20- and 40-day moving averages at the top with 20- and 40-day momentum along the bottom. The 40-day momentum (dark line) is smoother and peaks at about the same place as the price. The faster 20-day momentum is not as smooth but peaks sooner than the 40-day momentum and actually leads the price movement. The peaks of the faster momentum represent the maximum price change over a 20-day period. Had prices continued higher at the same rate—that is, if prices had increased $1 per day—the 20-day momentum would have turned sideways and become a horizontal line.

FIGURE 9.3 20- and 40-day momentum (lower panel) compared to a 20- and 40-day moving average, applied to AOL.

The 20-day and 40-day momentum lines both peak significantly sooner than their moving average equivalents, showing that the speed represented by momentum is a more sensitive measure than the direction of a trendline. The cost of this leading indicator is the increased noise seen in the erratic pattern of momentum compared to the smoothness of the trendline.

Momentum can be used as a trend indicator by selling when the momentum value crosses downward through the horizontal line at zero and buying when it crosses above the zero line. Because there is more noise in momentum than in the equivalent moving average, you would want to draw a small band around the zero line about the width of the small variations in momentum. A sell signal would be given when the momentum falls below the lower band and a buy when it moves upward through the upper band. However, the momentum crossing zero is essentially the same as the moving average turning up or down. It means that the price today is the same as the price *n*-days ago. Chapter 8 compared the performance of a trend and a momentum system over the same calculation periods.

Momentum as a Percentage

It is always convenient to express all markets in the same notation. It makes comparisons much easier. For the stock market, momentum can be expressed as a percentage where a 1-day momentum is equivalent to the 1-day return,

$$1\text{-day momentum in percent} = \frac{(p_t - p_{t-1})}{p_{t-1}} = \frac{p_t}{p_{t-1}} - 1$$

or, because they are returns, the alternative can be used, $\ln \dfrac{p_{t-1}}{p_t}$. But momentum is most often calculated over more than one day; therefore, the n-day momentum as a percentage is

$$M(n) = \frac{p_t}{p_{t-n}} - 1$$

Using percentages does not work for futures markets because most data used for analysis are continuous, back-adjusted prices. Back-adjusting over many years and many contracts will cause the oldest data to be very different from the actual prices that occurred on those dates; therefore, percentages are incorrect. When using momentum with back-adjusted futures prices, it is best to use the price differences.

Momentum as the Difference between Price and Trend

The term *momentum* is very flexible. It is common for it to refer to the difference between today's price and a corresponding moving average value, in a manner similar to *beta*, which indicates the relative strength of a stock compared to an index. The properties of this new value are the same as standard momentum. As the momentum becomes larger, prices are moving away from the moving average. As it moves towards zero, the prices are converging with the moving average.

The upper panel of Figure 9.4 shows a 20-period moving average plotted on a daily chart of Intel during the 14 months ending May 2003. The center panel shows the standard 20-period momentum, the difference between prices that are 20 days apart. The bottom panel is the difference between the price and the 20-day moving average. The range of values for the center panel is approximately +6 to −11, and the scale on the lower panel is +3 to −7. Because a trendline lags behind price movement, the difference between the price and trendline is smaller at points where there are larger price swings and typically larger momentum values.

A comparison of the two lower panels of Figure 9.4 also shows that there are differences in the small movements but the overall pattern is the same. When a price moves to an extreme, it is generally farthest from the moving average and farthest from the price n-days ago; therefore, the most extreme points occur at exactly the same place. This momentum calculation has also been called *relative strength* because it is measured relative to a previous price or relative to a trendline.

Momentum as a Trend Indicator

The momentum value is a smoothing of price changes and can serve the same purpose as a standard moving average. Many applications use momentum as a substitute for a price trend. By looking at the net change in price over the number of days designated by an n-day momentum indicator, intermediate fluctuations are ignored, and the pattern in

FIGURE 9.4 Momentum is also called *relative strength*, the difference between two prices or a price and a moving average (lower panel). The traditional momentum calculation is shown in the center panel.

price trend can be seen. The longer the span between the observed points, the momentum calculation period, the smoother the results. This is very similar to faster and slower moving averages.

To use momentum as a trend indicator, choose any calculation period. A *buy* signal occurs whenever the value of the momentum turns from negative to positive, and a *sell* signal is when the opposite occurs, as shown in Figure 9.5. If a band is used to establish a neutral position or a commitment zone, as discussed in the previous chapter, it should be drawn around the horizontal line representing the zero momentum value.

FIGURE 9.5 Trend signals using momentum.
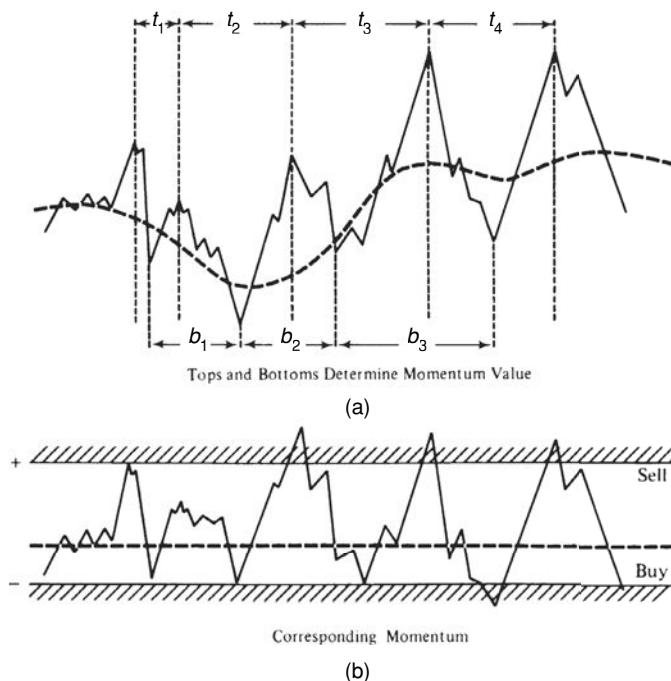
FIGURE 9.6 Relationship of momentum to prices. (a) Tops and bottoms determine momentum value. (b) Corresponding momentum.

In order to find the best choice of a momentum span, a sampling of different values could be tested for optimum performance, or a chart could be examined for some natural price cycle. Identify the significant tops and bottoms of any bar chart, and average the number of days between these cycles, or find the number of days that would closely approximate the occurrences of these peaks and valleys. Then use ½ or ¼ the number of days from peak to peak or valley to valley. These natural cycles will often be the best choice of momentum calculation interval (Figure 9.6). Momentum and oscillators, however, are more often used to identify abnormal price movements and for timing of entries and exits in conjunction with a longer-term trend.

Timing an Entry

Momentum is a convenient way of identifying good entry points based on small price reversals within a trend that are not large enough to change the direction of the trend. By choosing a much shorter time period for the momentum calculation (for example, 6 days) to work in combination with a longer trend of 30 to 50 days, the momentum indicator will show frequent opportunities within the trend. The short time period for the momentum calculation assures you that there will be an entry opportunity within about three days of the entry signal; therefore, it becomes a practical timing tool. This is particularly useful

when you view momentum as a countertrend signal, which will be discussed in the next section. Examples of using a momentum as entry timing can also be found in Chapter 19.

Identifying and Fading Price Extremes

An equally popular and more interesting use for momentum focuses on the analysis of relative tops and bottoms. All momentum values are bounded (although somewhat irregularly) in both directions by the maximum move possible during the time interval represented by the span of the momentum. The conditions at the points of high positive and negative momentum are called *overbought* and *oversold*, respectively.

A market is *overbought* when it can no longer sustain the strength of the current upwards trend and a downward price reaction is imminent; an *oversold* market is ready for an upward move. Faster momentum calculations (those using shorter calculation periods) will tend to fluctuate above and below the zero line based on small price changes. Longer calculation periods will take on the characteristics of a trend and stay above or below the zero line for the extent of the trend. In Figure 9.7, using S&P futures, the 20-day momentum in the center panel stays above zero for most of the period from September 2010 through February 2011. The very fast 3-day momentum in the bottom panel consistently penetrates the zero line, although the peaks on the upside are larger than those on the downside.

The values on the right scale of the two lower panels in Figure 9.7 show the range of price movement over the 20-day and 3-day periods. The 20-day scale goes from +100 to –60 while the 3-day scale is smaller, about +60 to –40. The 3-day momentum is not 3/20ths of the 20-day momentum because volatility does not increase linearly over time. That is, volatility can increase faster in a few days, but over longer periods it has up and down movement that limits the total move in one direction.

A system can take advantage of the momentum extremes by *fading the price movement* (selling rallies and buying declines). This is done by drawing two horizontal lines on the
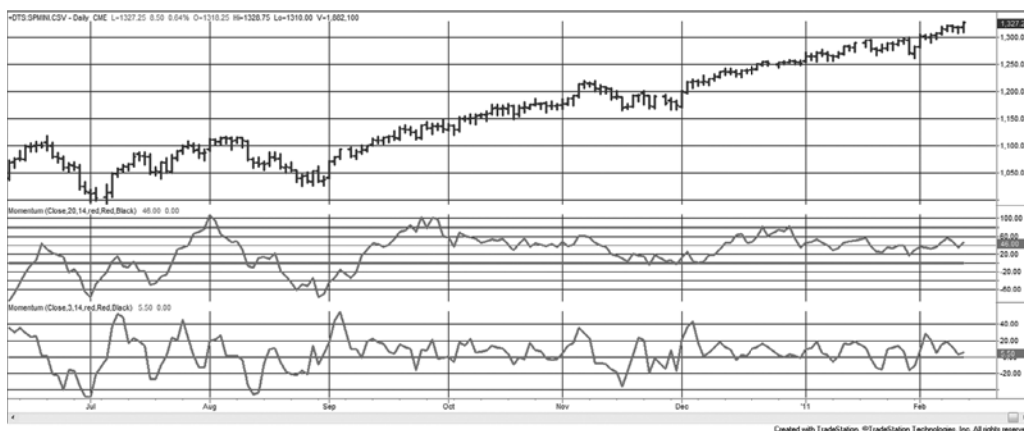
FIGURE 9.7 20-day momentum (center panel) and 3-day momentum (bottom panel) applied to the S&P, June 2010 through February 2011.

momentum chart (as shown in Figure 9.6b) above and below the zero line in such a way that the tops and bottoms of the major moves are isolated. These lines may be selected

- Visually, so that once the line is penetrated, prices reverse shortly afterwards.
- Based on a percentage of the maximum possible momentum value.
- A multiple of the standard deviation of momentum values. Using two standard deviations would position the lines so that only 5% of all momentum values penetrate above and below the two horizontal lines.

When positioning these bands, there is always a trade-off between finding more trading opportunities and entering the market too soon. This can be a complicated choice and is discussed in "System Trade-Offs" in Chapter 22. Once these lines have been drawn, the entry options will be one of the following basic trading rules:

- *Aggressive.* Enter a new long position when the momentum value penetrates the lower bound; enter a new short position when the value penetrates the upper bound.
- *Minor confirmation.* Enter a new short position on the first day the momentum value turns down after penetrating the upper bound (the opposite for longs).
- *Major confirmation.* Enter a new short position when the momentum value penetrates the upper bound moving lower (the opposite for longs).
- *Timing.* Enter a new short position after the momentum value has remained above the upper bound for *n* days (the opposite for longs).

To close out a profitable short or long position, there are the following alternatives (the rules are symmetrical for longs and shorts):

- *Most demanding.* Close out long positions or cover shorts when the momentum value satisfies the entry condition for a reverse position.
- *Moderately demanding.* Cover a short position when the momentum penetrates the zero line, minus one standard deviation (or some other target point partway between the zero line and the lower bound).
- *Basic exit.* Cover a short position when the momentum value becomes zero.
- *Allowing an extended move.* Cover a short position if the momentum crosses the zero line moving up *after* penetrating that line moving down.

The *basic exit*, removing your trade when the momentum reaches the zero line, is the benchmark case. It is based on the conservative assumption that prices will return to the center of the recent movement (in other words, *mean reverting*). It is less reasonable to assume that prices will continually move from overbought to oversold and back again. Most statisticians would agree that the mean is the best forecast of future prices, and the zero momentum value represents the mean. You may try to increase the profits by assuming that the momentum value will penetrate the zero line by a small amount because there is always some amount of noise. However, if you require a larger penetration
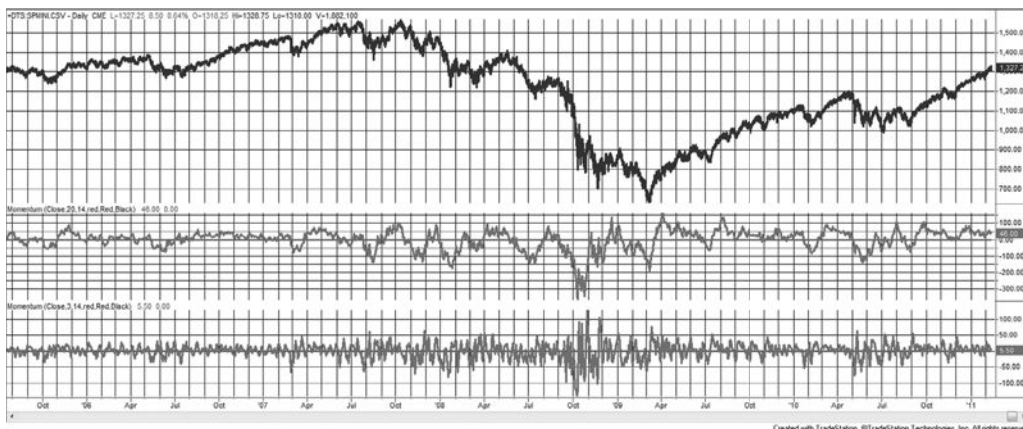
FIGURE 9.8 A longer view of the 20- and 3-day momentum applied to the S&P. The magnitude of the momentum values change over time.

of zero to cause an exit, many of the exit opportunities will be lost. You can be more conservative by exiting somewhat before the zero line is reached, but then you also reduce your average profit.

The same momentum calculations, viewed over a longer time period, show one of the problems with using momentum for timing. Figure 9.8 covers a critical period, from August 2005 through February 2011. Notice that the scale of the momentum values change dramatically from the beginning to the end of the chart. This is especially important for the 20-day momentum in the middle panel. For all of 2006, volatility was very low. Had we used momentum as an overbought/oversold indicator, the entry bands would have been at about ±30. For the 3-day momentum, it would have been closer to ±15. As we saw before, the faster momentum is more symmetric than the slower one. During 2006, there were more selling than buying opportunities.

Volatility increases in 2007. If buy entries were targeted as momentum values of −30 then traders would have held a 120 point loss as momentum dropped to −150. A similar pattern was repeated on the upside for sell signals. It would be necessary to increase the width of the buy and sell bands to ±75 and still hold larger losses waiting for a price reversal. But then comes 2008, and any countertrend entry, buying as prices fall, would have caused a complete loss of equity. However, let's say that we were able to avoid disaster and the entry-exit bands were at ±100 for the 20-day momentum and ±50 for the 3-day. Volatility finally declines and we get only the occasional entry signal. The bands need to be adjusted back down to a smaller range.

The problem with using momentum for timing is that the scale is unpredictable. Even if we waited for a reversal after penetrating the band there is no assurance that prices will not reverse and go to new extremes. Some risk protection is needed, a way of dynamically adjusting the bands (as discussed in Chapter 8), or a different way of measuring momentum.

Risk Protection

A protective *stop-loss order* can be used whenever you trade against the current price momentum. A stop-loss is a specific order to exit if the price goes the wrong way (discussed in Chapter 23 but also throughout the book). This is most important with the *aggressive* entry, which advocates selling when the momentum value moves above the upper threshold line, regardless of how fast prices are rising. While the other entry options allow for a logic positioning of a stop-loss, the aggressive entry does not. You will simply need to decide, based on historic momentum values, where to limit your risk. With the other entry options, prices are no longer at their extremes, so that stops may be placed using one of the following techniques:

- Place a protective stop above or below the most extreme high or low momentum value.
- Follow a profitable move with a trailing, nonretreating stop based on fixed points or a percentage.
- Establish zones that act as levels of attainment (using horizontal lines of equal spacing), and do not permit reverse penetrations once a new zone is entered.

Risk protection must be flexible in the way it deals with volatility. As prices reach higher levels, increased volatility will cause momentum tops and bottoms to widen; with low volatility, prices may not be active enough to penetrate the upper or lower bounds. This is shown in Figure 9.8. Therefore, both the profits and risk of a trade entered at extremes must increase with volatility and higher prices. Stops could be based on a percentage of price or a multiple of volatility.

The aggressive entry option remains the greatest problem for controlling risk, where a short signal is given when prices are very strong. If an immediate reversal does not occur, large open losses may accrue. To determine the proper stop-loss amount, a computer test was performed using an immediate entry based on penetration of an extreme boundary and with a stop-loss at the close of the day for risk protection. The first results were thought to have outstanding profits and consistency until it was discovered that the computer had done exactly the opposite of what was intended—it *bought* when the momentum crossed the upper bound and placed a close stop-loss *below* the entry. It did prove, for a good sampling of futures markets, that *high momentum* periods continued for enough time to capture small but consistent profits. It also showed that the aggressive entry trading rule, anticipating an early reversal, would be a losing strategy. One should never forget that declining momentum does not mean that prices are falling.

The Trend Provides Protection

In these examples, momentum is not a strategy itself, but a way of entering a trend trade. When a trend signal first occurs, it is not entered. If the signal was long, then the strategy waits until the 3-day momentum penetrates the lower band. If prices continue lower, then the trend will change from up to down, and the position will be exited. It would not be necessary to have a stop-loss.

Profit Targets for Fading Extreme Moves

When using momentum as a mean-reversion strategy, the most reasonable exit for a short entered at an extreme momentum high, or a long position entered at a low, is when the momentum returns to near zero. Prices do not have an obligation to go from extreme lows to extreme highs; however, any extreme can be expected to return to normal, which is where momentum is zero. Returning to zero does not mean that all trades will be profitable. If there has been a strong upwards move that lasted longer than the calculation period of the momentum, then prices will be higher when momentum finally returns to zero. The trend is not your friend if it fights with your mean-reversion position.

The number of profitable trades can be increased by targeting a momentum level that is more conservative than zero. That is, if a long was entered at an extreme low, a sell would be placed 5% or 10% below the zero momentum line, based on the distance between the buying threshold line and zero. Trades that *almost* return to normal would then be closed out with profits, although the average size of the profit would be smaller. To increase the size of the profits per trade, you would target the opposite side of the zero momentum line; therefore, if you held a long position, you would wait until the momentum value was greater than zero by 5% or 10% of the range. The success of taking profits early or late depends on the amount of noise in the momentum values as it fluctuates near the zero value. In many cases where extreme prices are followed by a sideways pattern, momentum will fluctuate around zero so that there is ample opportunity to exit the trade at a better price.

High-Momentum Trading

Some professional traders have made a business of trading in the direction of the price move only when momentum exceeds the *high threshold level* rather than anticipating a change of direction. There is a small window of opportunity when prices are moving fast. They will usually continue at the same, or greater, momentum for a short time, measured usually in minutes or hours, but occasionally a few days. There is a lot of money to be made in a short time—but at great risk if prices reverse direction sooner than expected.

Price patterns have changed. There are many more day traders. When one stock breaks out above a previous high, everyone sees it as an opportunity for profit. Buy orders start to flow, and volume increases. Stocks that are normally ignored can attract large volume when prices make a new high. Traders ride the rising prices for as long as possible, watching to see when volume begins to drop, then they exit. Some may just target a modest profit and get out when that target is reached.

High-momentum trading is a fast game that requires tools that allow you to scan a wide range of stocks. You are looking for one that has made a new high after a long, quiet period. You may also continually sort stocks by the highest momentum values. You need to stay glued to your screen, enter fast and exit fast. It is a full-time commitment.

Moving Average Convergence/Divergence (MACD)

Many of the practical problems of fading prices using momentum are solved with the *Moving Average Convergence/Divergence* (MACD),[2] developed by Gerald Appel. The MACD uses momentum calculated as the difference between two trendlines, produced using exponential smoothing. This momentum value is further smoothed to give a *signal line*. The most common form of MACD uses the difference between a 12-day and 26-day exponential smoothing. The signal line, used to produce trading recommendations, is a 9-day smoothing of the MACD. The MACD can be created as follows:

Step 1: Choose the two calculation periods for the trend, for example, 20 and 40. Convert to smoothing constants using $2/(n + 1)$.

Step 2: Calculate the slow trendline, $E40$, using the smoothing value 0.0243.

Step 3: Calculate the fast trendline, $E20$, using the smoothing value 0.0476.

Step 4: The *MACD line*, the slower-responding line in the bottom panel of Figure 9.9, is $E20 - E40$. When the market is moving up quickly, the fast smoothing will be above the slow, and the difference will be positive. This is done so that the MACD line goes up when prices go up.

Step 5: The *signal line* is the 9-day smoothing (using a constant of 0.10) of the MACD line. The signal line is slower than the MACD; therefore, it can be seen in the bottom panel of Figure 9.9 as the lower line when prices are moving higher.

The 20- and 40-day momentum lines, corresponding to the MACD line and the histogram, are shown in the center panel of Figure 9.9. It is clear that the MACD process smoothes these values significantly; therefore, it makes trading signals easier to see.

Reading the MACD Indicator

The MACD is normally seen as it appears at the bottom of Figure 9.9. The MACD line is higher in an uptrend and lower in a downtrend. The histogram is created by subtracting the slower signal line from the MACD line. When the histogram is above zero, it confirms the uptrend.

The 40/20 MACD line is similar to the 40-day momentum line. We can compare the MACD in the lower panel with the 20- and 40-day momentum in the center panel and see that the peaks and valleys are in about the same place, but the MACD line is much smoother because it is the difference between two trendlines rather than prices.

You should note the point at which the MACD and signal lines turn down in May 2001, compared to the corresponding moving averages that are shown with prices in the top panel of the chart. The MACD leads the moving averages in showing the downturn and has only slightly more up-and-down variation in its values. The signal line also leads the moving averages in the downturn and is equally smooth.

---

[2]From John D. Becker, "Value of Oscillators in Determining Price Action," *Futures* (May 1994).

FIGURE 9.9   MACD for AOL. The *MACD line* is the faster of the two trendlines in the bottom panel; the *signal line* is the slower. The histogram is created by subtracting the slower signal line from the MACD line.

Trading the MACD

The most common use of the MACD is as a trend indicator. For this purpose, only the MACD and signal lines are used in the following way:

> *Buy* when the MACD line (faster) crosses upwards through the signal line (slower).
> *Sell* when the MACD line crosses from above to below the signal line.

In the bottom panel of Figure 9.9, the buy signals that occur right after an extreme low in April and October of 2001 generated large gains. Unfortunately, there were many other crossings that generated losses; therefore, it is necessary to select which trades to enter. To accomplish that, the MACD uses thresholds similar to those shown for momentum, but in the opposite way. The MACD must first penetrate the lower band; then it must signal a new uptrend before a long position can be entered. In this way, it removes some of the whipsaws that might occur in a sideways market. Threshold levels were established by observing the historically high and low momentum values, shown as horizontal lines at the +2.00 and −2.00 levels in Figure 9.9. Sell signals would have been taken only after the MACD value had been above +2.00, and buy signals only when the MACD value had fallen below −2.00. The trading signals that satisfy these conditions are the *buy* signal in April, a *sell* in May, and a *buy* in October. However, fitting the threshold lines

make them unreliable values for trading. One solution can be found in "An RSI Version of MACD" in the next section.

Appel has written extensively on the MACD, and many variations can be found in his book, *Technical Analysis.*[3] Many of the examples are interpretive, similar to the way we look at chart patterns. Appel also preferred using trends of 19 and 39 days for the NASDAQ composite. An equally important use of MACD is for divergence signals. *Bearish divergence* occurs when prices are rising but the MACD values are falling. Because this technique applies to many indicators in addition to the MACD, it is covered in detail in the section "Momentum Divergence" toward the end of this chapter.

DIVERGENCE INDEX

A method similar to MACD but one that uses an interesting combination of generalized techniques is the *divergence index*, the volatility-adjusted difference between two moving averages, for example, 10 and 40 days. Using general notation:

1. *Slow average$_t$ = 40-day average of the most recent prices*
2. *Fast average$_t$ = 10-day average of the most recent prices*
3. *Difference$_t$ = price$_t$ – price$_{t-1}$*

Then the divergence index (DI),

$$DI_t = \frac{Fast\ average_t - Slow\ average_t}{StDev(Difference_t, slow\ period)^2}$$

The standard deviation of the price differences, taken over the slow period of 40, volatility-adjusts the results. The trading rules require a band around zero to trigger entries.

$$Band_t = stdev(DI_t, slow\ period)$$

$$Upper\ band_t = factor \times Band_t$$

$$Lower\ band_t = -factor \times Band_t$$

where factor = 1.0. By using the standard deviation of DI, the band also adjusts to changes in volatility. In Figure 9.10 the index in the lower panel has high peaks followed by periods of low or varying volatility. The standard deviation allows the bands to widen or narrow according to the pattern of fluctuations. This example used the slow period for the standard deviation; however, the fast period would have caused more rapid changes to the band.

---

[3] Gerald Appel, *Technical Analysis, Power Tools for Active Investors* (Upper Saddle River, NJ: FT Prentice Hall, 2005).
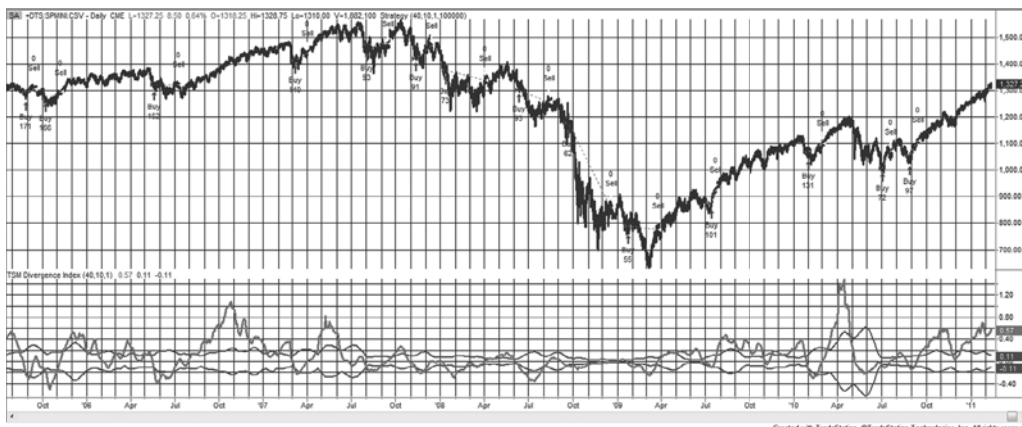
FIGURE 9.10 Divergence index applied to S&P futures.

The trading rules are

Buy when the DI moves below the lower band while in an uptrend
Sell when the DI moves above the upper band while in a downtrend
Exit longs and shorts when the DI crosses zero

Figure 9.10 shows the DI and the upper and lower bands in the bottom panel. The top panel shows the trading signals for the S&P 500 futures contract.

OSCILLATORS

Because the representation of momentum is that of a line fluctuating above and below a zero value, it has often been termed an *oscillator*. Even though it does oscillate, the use of this word is confusing. In this presentation, the term *oscillator* will be restricted to a specific form of momentum that is *normalized* and expressed in terms of values that are limited to the ranges between +1 and –1 (or +100 to –100), +1 and 0, or +100 and 0 (as in a percent).

To transform a standard momentum calculation into the normalized form with a maximum value of +1 and a minimum value of –1, divide the momentum calculation by its maximum value over the same rolling time period. This allows the oscillator to self-adjust to changes in price or volatility. For example, in April 2006, Amazon was trading at about $34. During a 10-day period it had a high of $35.31 and a low of $31.52, a range of $3.79. During 10 days in February 2011, Amazon had a high of $191.40 and a low of $174.77, a range of $16.63. Therefore, a price move of $2 in 2006 would be the same as a move of $8.77 in 2011. By dividing $2 by $3.79 or $8.77 by $16.63 we get 0.52 in both cases. Hence, *normalization*, which is the basis for most momentum indicators, can be an excellent way to self-adjust to price changes.

The following sections show how a number of useful oscillators are calculated. In each case, the purpose is to have the indicator post high values when prices are at a peak and low values when they are in a valley. To be most useful, prices should reverse direction soon after the oscillator records values near its extremes.

Relative Strength Index

One of the most popular indicators for showing overbought and oversold conditions is the *Relative Strength Index* (RSI) developed by Welles Wilder.[4] It provides added value to the concept of momentum by scaling all values between 0 and 100. It is more stable than momentum because it uses all the values in the calculation period rather than just the first and last. It is a simple measurement that expresses the relative strength of the current price movement as increasing from 0 to 100. It is calculated as

$$RSI = 100 - \left(\frac{100}{1 + RS} \right) = 100 \times \left(\frac{RS}{1 + RS} \right)$$

where $RS = \dfrac{AU}{AD}$

AU = the total of the upwards price changes during the past 14 days
AD = the total of the downwards price changes (used as positive numbers) during the past 14 days

Once the first calculation has been made, both the AU and AD values can be calculated daily using an *average-off* method:

$$AU_t = AU_{t-1} + \frac{AU_{t-1}}{14} + \max(p_t - p_{t-1}, 0)$$

$$AD_t = AD_{t-1} + \frac{AD_{t-1}}{14} + \max(p_{t-1} - p_t, 0)$$

This method essentially subtracts an average value and adds the new value, if any (see "Average-Off Method" in Chapter 7). All price changes are treated as positive numbers. The daily calculation of the RSI becomes a matter of simple arithmetic. Wilder has favored the use of 14 days because it represents one-half of a natural cycle, in this case, 1 month. He has set the significant threshold levels for the RSI at 30 and 70. Penetration of the lower level is indicative of an imminent upturn and penetration of the upper level, a pending downturn. A chart of an RSI is shown along with a comparable momentum and stochastic indicator in the next section, Figure 9.15.

---

[4]J. Welles Wilder Jr., *New Concepts in Technical Trading Systems* (Greensboro, NC: Trend Research, 1978).

FIGURE 9.11 RSI top formation.

Use of the RSI alone to generate trading signals often requires interpretation similar to standard chart analysis. Lines are drawn across the tops of the RSI values to indicate a downtrend. A head-and-shoulders formation can be used as the primary confirmation of a change in direction. Wilder himself used the RSI top and bottom formations shown in Figure 9.11. A break of the reaction bottom between the declining tops is a sell signal. In addition, the *failure swing*, or divergence (discussed later in this chapter), denotes an unsuccessful test of a recent high or low RSI value.

Wilder created other popular indicators. One of these is a momentum calculation called the *Average Directional Movement* (ADX). The ADX is a byproduct of the *Directional Movement* and is discussed in Chapter 23. All of these indicators are actively used by traders.

Modifying the RSI

An obvious objection to the RSI might be the selection of a 14-day *half-cycle*. Maximum divergence is achieved by using a moving average that is some fraction of the length of the dominant cycle, but 14 days may not be that value. If a 14-day calculation period is too short, then the RSI would remain outside the 70-30 zones for extended periods rather than signaling an immediate turn. In practical terms, a 14-day RSI means that a sustained move in one direction that lasts for more than 14 days will produce a very high RSI value. If prices continue higher for more than 14 days, then the RSI, as with other oscillators, will go sideways. The idea is to pick the calculation period for which there are very few larger sustained moves, and 14 may be that value. If more frequent overbought and oversold conditions are needed, then the period could be lowered to 10. At the same time, the zones could be increased to 80-20. Some combination of calculation period and zone will usually give the frequency of trades that is needed.

A study by Aan[5] on the distribution of the 14-day RSI showed that the average RSI top and bottom value consistently grouped near 72 and 32, respectively. Therefore, 50% of all RSI values fall between 72 and 32, which can be interpreted as *normally distributed*, and

---

[5] Peter W. Aan, "How RSI Behaves," *Futures* (January 1985).

equivalent to about 0.675 standard deviations. This would suggest that the 70-30 levels proposed by Wilder are too close together to act as selective overbought/oversold values, but should be moved farther apart. The equivalent of 1.5 standard deviations is a comfortable trade-off between frequency of trades and risk. It is generally safer to err on the side of less risk. If there are too many trades being generated by the RSI, a combination of a longer interval and higher confidence bands will be an improvement.[6]

Further Smoothing with *N*-Day Ups and Downs

Instead of increasing the number of days in the RSI calculation period, a smoother indicator can be found by increasing the period over which each of the *up* and *down* values are determined. The original RSI method uses 14 individual days, where an *up day* is a day in which the price change was positive. Instead, we can replace each 1-day change with a 2-day change, or an *n*-day change. If we use 2-day changes, then a total of 28 days will be needed, so that each 2-day period does not overlap another; there will be 14 sets of two days each. Using 14 sets of 2 will give a smoother indicator than using 28 single-day changes.

RSI Countertrend Trading

Momentum indicators are used for timing and countertrend trading. Normally, it is the absolute overbought or oversold level that is the trigger for a trade. However, either a very fast move in price or an extreme value for a momentum indicator could be a criterion for a reversal. When $RSI_t - RSI_{t-k} > a$, or when $RSI_t - RSI_{t-k} < -a$, where a is a threshold value, we could sell and buy, respectively. This also increases the number of trades because it does not require that $RSI_t > 0$, only that it has moved quickly. Exits can be taken when momentum returns to near zero, or after n days. A program to test this is *TSM RSI Countertrend*, available on the Companion Website.

Net Momentum Oscillator

Another variation on the RSI is the use of the difference between the sum of the up days and the sum of the down days, called a *net momentum oscillator*.[7] If you consider the unsmoothed RSI = $100 \times (S_u/(S_u + S_d))$ then the net momentum oscillator would be

$$CMO = 100 \times (S_u - S_d)/(S_u + S_d)$$

This method replaces some of the indicator movement lost to smoothing in the normal RSI, and shows more extremes. This may also be done by shortening the number of periods in the RSI calculation.

---

[6]For another interesting approach to RSI optimization, see John F. Ehlers, "Optimizing RSI with Cycles," *Technical Analysis of Stocks & Commodities* (February 1986).

[7]Tushar S. Chande and Stanley Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994).

The 2-Day RSI

Among the interesting information published on MarketSci Blog[8] is Michael Stokes's "favorite" indicator, the 2-day RSI, which replaces each 1-day closing price change with a 2-day change; otherwise, the calculations and the number of data points are the same. If you already have a program to calculate the RSI, you only need to change the statement $p_t - p_{t-1}$ with $p_t - p_{t-2}$ to get overlapping 2-day data. Or, to try an n-day RSI, replace the "2" with your choice of days. The effect of using combined 2 days instead of 1 is some smoothing and a general increase in volatility.

Stokes uses threshold levels of 10 and 90 for the S&P and the rules

Buy on the next close when the 2-day RSI penetrates the threshold of 10 moving lower

Sell short on the next close when the 2-day RSI penetrates the threshold of 90 moving higher

Exit one day later on the close

Figures 9.12a and b are from the MarketSci Blog on December 9, 2008. Figure 9.12a shows that the profitability of the RSI reversed in about 1998. From 1970 to 1998 it was a good trend indicator; that is, when the RSI moved above 90 the S&P continued up. After 1998 it has been a much better mean-reverting indicator. Figure 9.12b is the result of applying a method of scaling in to trades according to the following rules:

If the RSI < then *buy* this %		If the RSI > then *sell short* this %	
< 5	100%	> 95	100%
< 10	75%	> 90	75%
< 15	50%	> 85	50%
< 20	25%	> 80	25%

On day t the RSI was 14, and we entered 50%, but then on day $t + 1$ the RSI dropped to 8; it was not clear whether we added an extra 25%, but we assume that is the case. Then the trade is exited if we do not add on the next day. Because the RSI can stay above or below 50 for long periods of time, an exit rule that waits for the RSI to cross 50 is not likely to be successful. Stokes's approach of exiting after one day seems much safer. Figure 9.12b separates the longs and short sales from 2000 through 2008, showing that both sides performed well, a very desirable outcome. Having good performance for longs and shorts is not common because many markets have an upwards bias (especially the stock market). However, 1-day trades are sensitive to costs. Traders will need to see if the returns per share, or per contract, are sufficient to cover commissions and slippage.

---

[8]Free subscriptions are available at MarketSci Blog email@marketsci.com. The author, Michael Stokes, presents his own ideas and analysis of various systematic approaches.

Evolution of RSI(2) Strategy (1970 to Present)

— RSI(2)<10 — RSI(2)>90
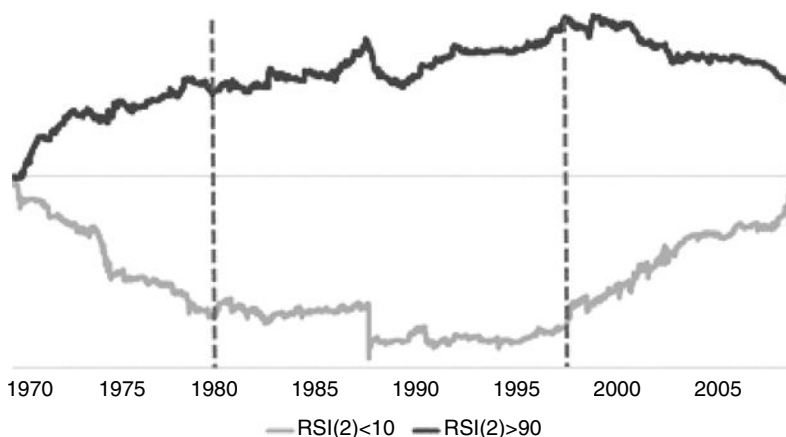
FIGURE 9.12a   Performance of the 2-day RSI, MarketSci blog's favorite oscillator, from the December 9, 2008, posting, simple entries <10 and >90 applied to the S&P, 1970 through 2008.

A Standard 2-Period RSI

In Figure 9.13 a standard 2-period RSI is shown in the middle panel and the traditional 14-day RSI in the lower panel. Both are applied to NASDAQ futures, from August 2005 through September 2008, which appears at the top. The classic RSI penetrates the extremes of 30 and 70 but does not reach 20 or 80, while the 2-period RSI touches near 15 and 85 a number of times. We might expect a 2-period oscillator to jump between 0 and 100 every day that prices changed direction, but the average-off calculation slows down the process and makes the 2-period chart a somewhat more volatile version of the standard RSI.

Trading Strategy: Long versus Short-Only

— Trading Strategy: Long Only   — Trading Strategy: Short Only

FIGURE 9.12b   A breakdown of longs and short sales using the scaling-in method, S&P 2000 through 2008.
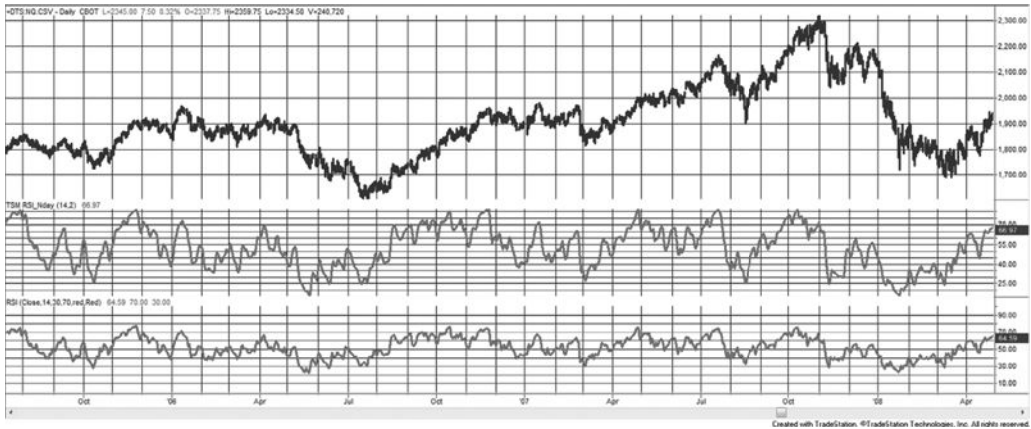
FIGURE 9.13 2-Period RSI (center panel) compared with the traditional 14-day RSI (bottom panel), applied to NASDAQ futures, August 2005 through September 2008.

An RSI Version of MACD

While the MACD creates a histogram of the difference between two moving averages, we can make that pattern easier to interpret by applying the RSI to the spread of the moving averages. In Figure 9.14 a 5-day RSI is applied to the difference between the 10- and 40-week moving average for the *e*mini-S&P futures. Using the standard 30-70 thresholds, this seems to find credible points where prices are overbought and oversold.
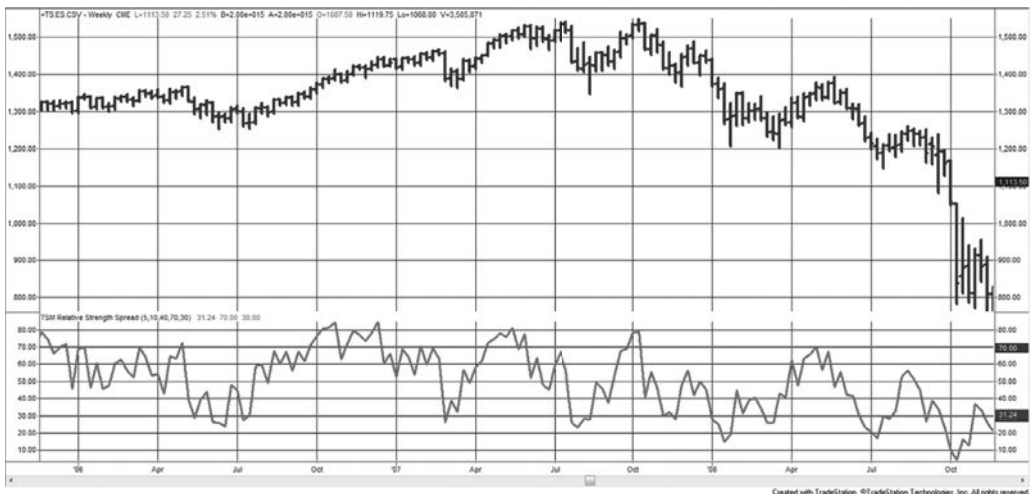
FIGURE 9.14 RSI applied to the spread between a 10-week and 40-week moving average of S&P futures.

Stochastics

The *stochastic* indicator, created by George Lane, is an oscillator that measures the relative position of the closing price within a past high-low range. It is based on the commonly accepted observation that closing prices tend to resist penetrating the high prices of the past few days, the place where a horizontal resistance line would be drawn on a chart. Similarly, in a downtrend prices must be able to close below the lows of the past few days for the move to continue. When the market is about to turn from up to down, for instance, it is often the case that the highs are higher than previous days, but the closing price settles nearer the low of the day, failing to indicate a continuation of the uptrend. This makes the stochastic oscillator different from the MACD, which uses the difference between two trends, and the RSI which uses only the closing prices. The stochastic uses the high, low, and close and unlike the other oscillators, there does not have to be any smoothing to introduce a lag.

The three indicators that result from the stochastic measurement are called *%K*, *%D*, and *%D-slow*. These indicators show increasingly slower interpretation of price movement, with *%D* being the most popular as a single indicator; however, *%D* and *%D-slow* are often used together to produce a trading signal. Calculation of these indicators for today's value t over the past 5 days, are

$$Initial\ (raw)\ \%K_t = 100 \times \frac{C_t - L_t(5)}{R_t(5)}$$

$$\%D = \text{``}\%K\text{-slow''} = \frac{\%K_t + \%K_{t-1} + \%K_{t-2}}{3} = \left(\frac{\sum_{i=t-2}^{t} \%K_i}{3}\right)$$

$$\%D_t\text{-slow} = \frac{\left(\sum_{i=t-2}^{t} \%D_i\right)}{3}$$

Where C_t = today's closing price

         $L_t(5)$ = the low price of the last 5 days

         $R_t(5)$ = the range of the last 5 days (highest high minus lowest low) as of today.[9]

Calculating the 10-Day Stochastic for Hewlett-Packard (HPQ)

Using an Excel spreadsheet, each of the stochastic components can be calculated in only a few columns. In Table 9.1, the historic prices for Hewlett-Packard are imported into columns A–D. The 10-day stochastic calculations, columns *E–J*, are:

  1. Column *E*, the 10-day high using the function *Max*.

  2. Column *F*, the 10-day low using the function *Min*.

---

[9]Harry Schirding, "Stochastic Oscillator," *Technical Analysis of Stocks & Commodities* (May/June 1984).

TABLE 9.1 Excel Example of 10-Day Stochastic for Hewlett-Packard (HPQ)

Row	Date	High	Low	Close	10-Day High	10-Day Low	10-Day Range	%K	%K-Slow	%D-Slow
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
1	1/3/2011	43.49	42.22	42.74						
2	1/4/2011	43.77	43.01	43.63						
3	1/5/2011	44.22	43.40	44.20						
4	1/6/2011	44.96	44.18	44.88						
5	1/7/2011	45.39	44.71	45.09						
6	1/10/2011	45.05	44.57	44.86						
7	1/11/2011	46.06	45.20	45.43						
8	1/12/2011	45.71	45.27	45.64						
9	1/13/2011	45.84	45.31	45.65						
10	1/14/2011	46.40	45.61	46.25	46.40	42.22	4.18	96.41		
11	1/18/2011	46.42	46.08	46.34	46.42	43.01	3.41	97.65		
12	1/19/2011	46.48	46.08	46.32	46.48	43.40	3.08	94.81	96.29	
13	1/20/2011	46.79	45.76	46.78	46.79	44.18	2.61	99.62	97.36	
14	1/21/2011	47.64	46.83	47.23	47.64	44.57	3.07	86.64	93.69	95.78
15	1/24/2011	47.59	46.66	47.55	47.64	44.57	3.07	97.07	94.44	95.16
16	1/25/2011	47.83	46.88	47.08	47.83	45.20	2.63	71.48	85.07	91.07
17	1/26/2011	47.28	46.57	46.88	47.83	45.27	2.56	62.89	77.15	85.55
18	1/27/2011	46.98	46.58	46.74	47.83	45.31	2.52	56.75	63.71	75.31
19	1/28/2011	46.69	45.36	45.51	47.83	45.36	2.47	6.07	41.90	60.92

3. Column *G*, calculate the high-low range by subtracting *F* from *E*.
4. Column *H*, row *n*, calculate the raw stochastic *%K* as (D*n*-F*n*)*100/G*n*.
5. Column *I*, the *%Kslow* is the 3-day average of *%K*, column *H*.
6. Column *J*, *%Dslow* is the 3-day average of *%K-slow*, column *I*.

The raw stochastic in column *H* of the spreadsheet shows that values start above 90% and drop to 6% in a few days. This happens when prices close near the highest or lowest prices of the past 10 days. When calculating the stochastic, today's close is always included in the *max*, *min*, and *range* calculations. If today's price is the highest or lowest of the calculation period, then the raw stochastic will have a value of 100 or 0. The full spreadsheet *TSM Stochastic calculation for HPQ*, can be found on the Companion Website.

Comparing the Stochastic to Momentum and the RSI

The calculations for momentum, RSI, and the stochastic are very different; therefore, we would expect a chart of these indicators to vary considerably from one another. However, there are surprising similarities among the three, even calculated over the same 14-day period, as shown in Figure 9.15. The momentum and stochastic both show how
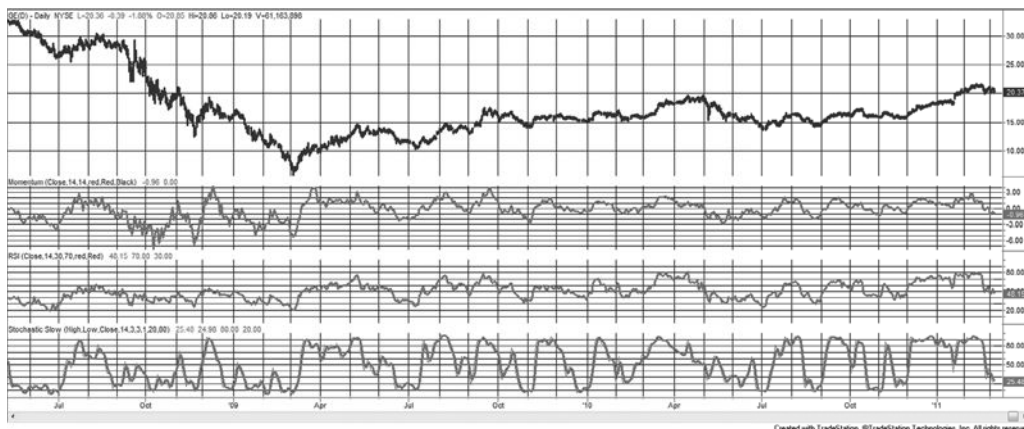
FIGURE 9.15   Comparison of simple momentum, RSI, and stochastic, all for 14-day calculation periods, General Electric, May 2008 through February 2011.

today's price relates to the prices over the calculation interval; therefore, we would expect them to be most similar, and that appears to be true. Both have peaks at the same place, although the *%D-Slow* stochastic has a smoother overall line due the two successive 3-day averages.

The RSI chart does not seem to be lagged, instead it is dampened, appearing to have less volatility. In fact, the highs and lows rarely touch 70 and 30, while the stochastic highs and lows often go above 90 and below 10, even though it is smoothed. The overall impression of the stochastic is that it moves faster and exaggerates the swings.

Trading the Stochastic

Traditionally, the stochastic can be traded by using a combination of a slower and faster calculation to give signals, or by combining extreme stochastic values for entry or exit timing with a trend. To use the stochastic by itself, *%D* and *%D-slow* can take on the role of the faster indicator and the slower signal line in the same way as the MACD and its signal line were used. The fastest calculation, *%K*, is not often used due to its instability.[10] Figure 9.16 shows the two 20-day calculations, *%D* and *%D-slow*, along the bottom of an S&P futures continuation chart from December 2001 through September 2002. A 60-day moving average is in the top panel.

The downtrend in the S&P during 2002 produces two good sell signals using the stochastics. In March and August, the slower *%D* penetrates above the 80% threshold and crosses the signal line moving down. A third peak in May touches near 70% and would produce an additional sell signal if the upper threshold were set lower. During a sustained downtrend, it is common for the indicator values to fall below the lower threshold

---

[10]It is common practice to use the notation %K and %D to mean %K-slow and %D-slow, respectively. All writings on the stochastic use the smoothed values, rather than the initial %K calculation, regardless of the omission of "-slow." Any use of %K in this text also refers to %K-slow unless specifically stated.
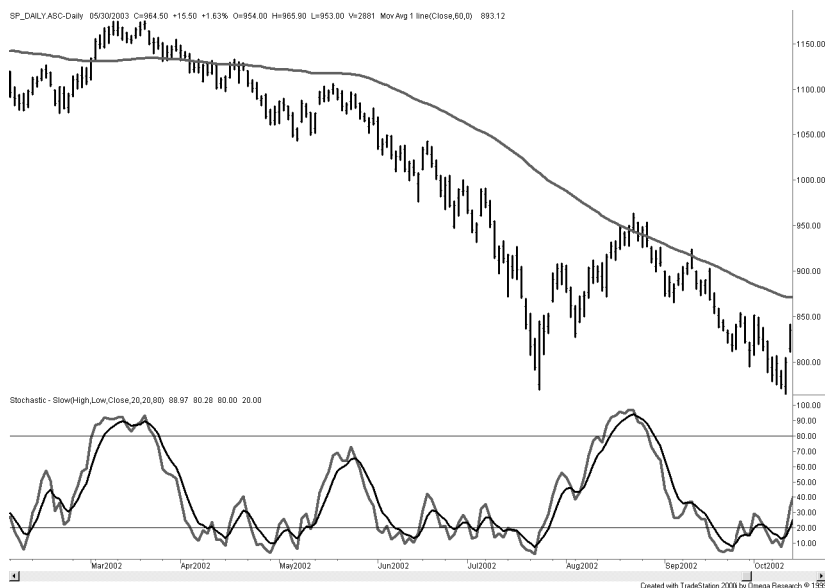
FIGURE 9.16 20-day stochastic (bottom) and a 60-day moving average for the S&P futures continuation series.

and remain there for extended periods. Although the rolling calculation is intended to self-adjust, the most practical solution is to combine the stochastic signals with a trend.

The 60-day moving average, shown along with prices in the upper part of Figure 9.16, indicates a downtrend for the entire period of the chart. Using the relatively slow 20-day stochastic, the penetration of the upper 80% threshold gives very good timing for short sale entries, and avoids the problems of an unfavorable distribution of stochastic values. The 20-day stochastic does not indicate where to exit the downtrend. A faster stochastic can be used to produce more frequent sell signals that can be applied as follows

> *Enter a short sale* after the trend has turned down on the first stochastic sell signal (the stochastic crosses the signal line going down).
> *Enter a long* after the trend has turned up on the first stochastic buy signal.

The problem with all timing rules is that, if prices move higher and do not retrace, then there will not be a stochastic signal, or it will occur well after there would have been profits in the trade. Many of the lost opportunities can be eliminated by using a faster stochastic calculation period. It is particularly dangerous to use a timing rule for exiting a position. Delays in entering are lost opportunities, but delays exiting are real trading losses.

Left and Right Crossovers

The faster *%K-slow* will usually change direction sooner than the *%D-slow*, crossing the *%D-slow* line while it is still moving in the prior trend direction. The opposite case, when
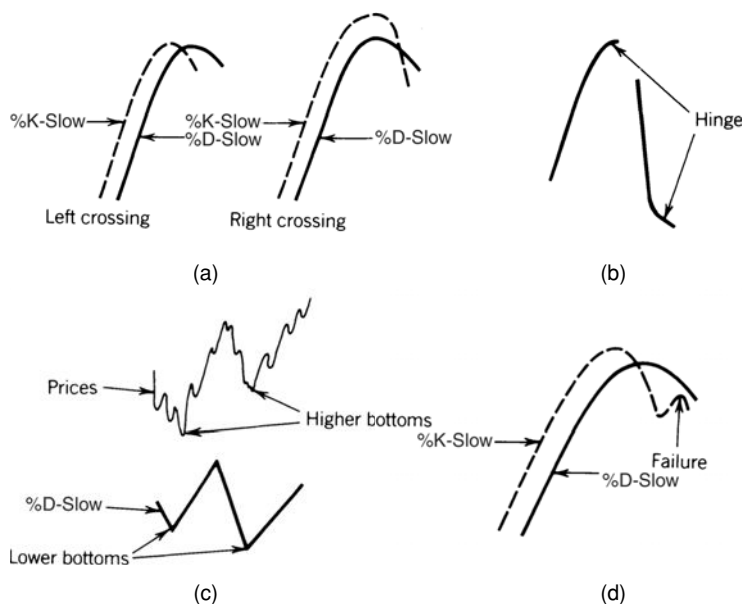
FIGURE 9.17 Lane's patterns. (a) Left and right crossings.
(b) Hinge. (c) Bear market set-up. (d) %K-slow failure.

the *%D-slow* turns first, indicates a slow, stable change of direction and is a more favorable pattern (Figure 9.17a). Using the fast *%K-Slow* and the slower *%D-Slow*, the following patterns have been identified:

> *Hinge.* A reduction in the speed of either the *%K-slow* or *%D-slow* lines, shown as a flattening out, indicates a reversal on the next day (Figure 9.17b).
>
> *Warning.* An extreme turn in the faster *%K-slow* (from 2 to 12%) indicates at most two days remaining in the old trend.
>
> *Extremes.* Reaching the extreme *%K-slow* values of 0 and 100 requires seven consecutive days of closes at the highs (or lows). The test of these extremes, following a pullback, is an excellent entry point.
>
> *Set-up.* Although the line chart shows higher highs and lows, if the *%D-slow* line has lower lows, a *bear market set-up* has occurred. Look for a selling opportunity on the next rally (Figure 9.17c).
>
> *Failure.* An excellent confirmation of a change in direction occurs when *%K-slow* crosses *%D-slow* (after penetrating the extreme level), then pulls back to the *%D-slow* line, but fails to cross it again (Figure 9.17d).

Creating a Stochastic from the RSI

Any series or indicator value can be converted to a raw stochastic, *%K*, without adding lag by replacing the closing price with the indicator value. This creates a measure of

where that indicator lies in its high-low range over the calculation period and may simplify the generation of buy and sell signals. A stochastic created from an RSI would be[11]

$$n\text{-}day\ StochRSI = \frac{RSI_{\text{today}} - RSI_{n\text{-day low}}}{RSI_{n\text{-day high}} - RSI_{n\text{-day low}}}$$

Williams's Oscillators

Larry Williams has been known for his development of trading methods based on oscillators since his publication of the *A/D oscillator* in 1972. The three techniques described below have some similarities, even though the last one, the *Ultimate Oscillator*, was published 13 years later.

A/D Oscillator

In 1972, Jim Waters and Larry Williams published a description of their *A/D Oscillator*. For their method, *A/D* means *accumulation/distribution* rather than the popular notation of advance/decline, a well-known indicator for stocks. They used a unique form of relative strength, defining *buying power* (BP) and *selling power* (SP) as

$$BP = high - open$$

$$SP = close - low$$

where the values used were today's open, high, low, and closing prices. The two values, BP and SP, show the additional buying strength (relative to the open) and selling strength (compared to the close) in an effort to measure the implied direction of the day's trading. This definition of buying and selling power is still used today. The combined measurement, called the *Daily Raw Figure* (DRF) is calculated as

$$DRF_t = \frac{BP_t + SP_t}{2 \times (H_t - L_t)}$$

The maximum value of 1 is reached when a market opens trading at the low and then closes at the high: $BP_t - SP_t = H_t - L_t$. When the opposite occurs and the market opens at the high and closes on the lows, the $DRF = 0$. Each price series develops its own patterns, which can be smoothed or traded in many ways similar to a momentum index. The Waters-Williams A/D Oscillator solves problems of volatility and limit moves (although there are very few markets with trading limits any more) in futures markets. *DRF* completely adjusts to higher or lower trading ranges because the divisor itself is a multiple of the day's range; because each day is treated independently, the cumulative values of the momentum index are not part of the results.

---

[11]Tushar Chande and Stanley Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994).

This day-to-day evaluation has no memory and causes DRF to be very volatile, yet it still takes on the pattern of the underlying market. If fewer trades are preferred, the DRF can be smoothed. In the example in Table 9.2, a 0.30 exponential smoothing was used (selected arbitrarily) and the entry bands narrowed from 80-20 to 70-30 to reflect fewer extreme prices. Because the 1973 examples used soybeans, this example will do the same, but for 2011 data. Table 9.2 shows the calculations for the DRF and the smoothed DRF. The DRF is plotted as bars on a scale of 0 to 1.00 in Figure 9.18 and is extremely erratic. The solid line is the smoothed DRF. Once plotted, two horizontal lines can be drawn to isolate the peaks and bottoms of DRF; the top part becomes a zone representing an overbought condition, and the bottom zone represents oversold. For this chart, the 80-20 levels were used, although we expect this to be consistent over time because the range used in the denominator normalizes volatility. Even though we do not claim that these thresholds are predictive, the original article by Waters and Williams had lines drawn in a similar place. The thresholds for the smoothed DRF were set at 70-30.

TABLE 9.2 A/D Oscillator and Trading Signals, Soybeans, January 1, 2011 through March 4, 2011

					Raw DRF			30% DRF Smoothing		
Date	Open	High	Low	Close	DRF	80-20 Signal	Entry (Lagged)	DRF	70-30 Signal	Entry (Lagged)
1/3/2011	1409.00	1422.00	1388.00	1392.00	0.250			0.250		
1/4/2011	1391.25	1403.50	1370.50	1382.50	0.367			0.285		
1/5/2011	1385.00	1408.00	1375.00	1406.50	0.826	Sell		0.447		
1/6/2011	1405.50	1411.75	1389.00	1391.00	0.181	Buy	1405.50	0.368		
1/7/2011	1389.75	1395.50	1373.25	1378.00	0.236	Buy	1389.75	0.328		
1/10/2011	1381.00	1407.50	1379.00	1393.50	0.719	Buy		0.445		
1/11/2011	1395.00	1401.50	1368.25	1370.00	0.124	Buy		0.349		
1/12/2011	1372.50	1440.00	1372.00	1428.00	0.908	Sell		0.517		
1/13/2011	1431.75	1445.50	1426.00	1429.00	0.429	Sell	1431.75	0.491		
1/14/2011	1431.50	1442.50	1418.00	1435.50	0.582	Sell		0.518		
1/18/2011	1425.25	1443.00	1421.25	1426.25	0.523	Sell		0.519		
1/19/2011	1429.50	1444.75	1415.25	1424.50	0.415	Sell		0.488		
1/20/2011	1426.00	1434.00	1396.50	1427.25	0.517	Sell		0.497		
1/21/2011	1428.75	1440.00	1418.00	1425.25	0.420	Sell		0.474		
1/24/2011	1425.50	1439.50	1409.75	1417.50	0.366	Sell		0.441		
1/25/2011	1417.25	1421.00	1383.25	1387.50	0.106	Buy		0.341		
1/26/2011	1388.50	1404.00	1377.25	1398.50	0.687	Buy	1388.50	0.445		
1/27/2011	1402.25	1414.00	1391.00	1412.50	0.723	Buy		0.528		
1/28/2011	1414.50	1436.25	1405.50	1411.00	0.443	Buy		0.503		
1/31/2011	1408.25	1426.50	1407.50	1426.00	0.967	Sell		0.642		
2/1/2011	1426.50	1453.50	1421.25	1451.00	0.880	Sell	1426.50	0.713	Sell	
2/2/2011	1452.50	1465.00	1446.50	1457.00	0.622	Sell		0.686	Sell	1452.50

TABLE 9.2 *(continued)*

Date	Open	High	Low	Close	Raw DRF			30% DRF Smoothing		
					DRF	80-20 Signal	Entry (Lagged)	DRF	70-30 Signal	Entry (Lagged)
2/3/2011	1459.00	1465.50	1447.00	1448.50	0.216	Sell		0.545	Sell	
2/4/2011	1451.50	1455.75	1435.25	1446.50	0.378	Sell		0.495	Sell	
2/7/2011	1451.00	1456.75	1434.25	1437.50	0.200	Sell		0.406	Sell	
2/8/2011	1438.00	1450.00	1427.00	1447.25	0.701	Sell		0.495	Sell	
2/9/2011	1450.00	1468.75	1448.00	1464.00	0.837	Sell		0.598	Sell	
2/10/2011	1464.25	1468.00	1443.00	1446.00	0.135	Buy		0.459	Sell	
2/11/2011	1444.75	1454.50	1423.00	1429.00	0.250	Buy	1444.75	0.396	Sell	
2/14/2011	1431.75	1438.00	1413.25	1416.00	0.182	Buy		0.332	Sell	
2/15/2011	1416.00	1422.00	1378.00	1381.25	0.105	Buy		0.264	Buy	
2/16/2011	1382.25	1389.00	1370.00	1378.50	0.401	Buy		0.305	Buy	1382.25
2/17/2011	1378.50	1418.75	1378.50	1416.50	0.972	Sell		0.505	Buy	
2/18/2011	1417.50	1421.25	1373.00	1381.00	0.122	Buy	1417.50	0.390	Buy	
2/22/2011	1391.00	1397.50	1311.00	1311.00	0.038	Buy	1391.00	0.284	Buy	
2/23/2011	1310.00	1335.00	1296.25	1331.50	0.777	Buy		0.432	Buy	
2/24/2011	1332.25	1336.75	1310.25	1329.25	0.443	Buy		0.436	Buy	
2/25/2011	1328.75	1399.25	1322.00	1375.00	0.799	Buy		0.545	Buy	
2/28/2011	1378.00	1380.25	1357.75	1364.75	0.206	Buy		0.443	Buy	
3/1/2011	1368.00	1376.50	1353.75	1375.25	0.659	Buy		0.508	Buy	
3/2/2011	1374.50	1399.00	1372.00	1394.25	0.866	Sell		0.615	Buy	
3/3/2011	1390.50	1414.50	1384.50	1412.00	0.858	Sell	1390.50	0.688	Buy	
3/4/2011	1410.25	1424.50	1395.75	1414.00	0.565	Sell		0.651	Buy	
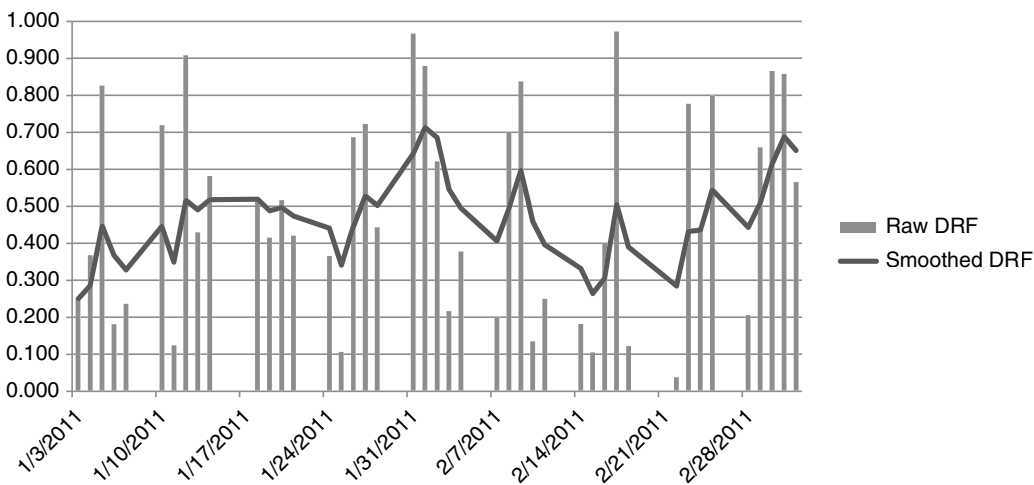
FIGURE 9.18 A/D Oscillator.

The rules for using the A/D Oscillator were not defined in the original article, but some simple rules could be:

- *Sell* when the DRF (or smoothed DRF) penetrates into the overbought zone. Close out all accumulated long positions (if any) and go short on the open of the next trading day.
- *Buy* and exit the short position when the opposite conditions occur.
- All positions are entered on the next open after a signal.

For risk control, positions could be exited if the DRF does not post a lower value (for shorts) within one or two days, and the opposite for longs. In Table 9.2 no stops were used. Long positions were held until a short signal occurred, and shorts until a long was signaled. If the DRF (or smoothed DRF) enters an overbought or oversold zone more than once without the opposite zone being entered, an additional position can be added, although that was not done in the example. Following these rules, the A/D Oscillator showed nine trades, six of them profitable. For the smoothed DRF there were only two trades, both profitable. The amount of smoothing can easily be changed. For more signals, increase the smoothing constant so that it's greater than 0.30. An Excel spreadsheet and a TradeStation program, both named *TSM AD Oscillator*, can be found on the Companion Website.

Although this example was very good, there are potential problems in the A/D Oscillator. The first happens on locked-limit days (where the market opens at the allowable limit and does not trade). In that case the open, high, low, and close are all the same and the DRF cannot be calculated because the divisor is zero. Fortunately, there are few markets in which that can happen, but it did recently in cotton. A more basic problem concerns gap openings. A much higher opening with a stronger close would also upset the resulting DRF value. Consider the following example:

	Open	High	Low	Close	DRF	ΔRFs	ΣΔDRF
Monday	43.00	44.00	40.00	41.00	0.25		
Tuesday	42.00	42.00	39.00	40.00	0.17	−0.17	−0.17
Wednesday	38.50	38.50	38.00	38.00	0.00	+0.17	0.00
Thursday	42.00	42.00	39.00	40.00	0.17	−0.33	−0.33
Friday	40.00	43.00	40.00	42.00	0.83	+0.50	+0.17

Note that on Wednesday, the DRF indicates that the momentum has reversed, but in fact the price is falling rapidly and gives no indication of recovering; it may actually be gaining momentum. On Thursday the price soars up and closes in the midrange, but the DRF shows a new downward momentum. The problem seems to be related to lack of association with the prior closing price. The daily movement can take on different appearances if the entire range was above or below the closing price. To form this link, replace the current high or low with the prior closing price, in the manner of the *true range* calculation, if that price was outside the current trading range. The following example shows that the results smoothed out and leaves the trend intact.

	Open	High	Low	Close	DRF	ΔDRF	ΣΔDRF
Monday	43.00	44.00	40.00	41.00	.25		
Tuesday	42.00	42.00	39.00	40.00	.17	−.17	−.17
Wednesday	38.50	(40.00)	38.00	38.00	.37	+.04	.13
Thursday	42.00	42.00	(38.00)	40.00	.25	−.12	−.25
Friday	40.00	43.00	40.00	42.00	.83	+.58	+.33

Linking the Current Day with the Prior Day

Another oscillator can be constructed using the highs and lows relative to the prior close:

$$O_t = \frac{H_t - C_{t-1}}{H_t - L_t}$$

The two days are linked together, and the ratio of the high price relative to the prior close is measured against the total range for the day. For the normal case, $H_t \geq C_{t-1} \geq L_t$; but if C_{t-1} replaces either H_t or L_t to extend the range, the value of O_t will be either 1 or 0 for these extreme cases. As with the A/D Oscillator, the values derived from this method may also be smoothed.

Oscillators are not the only tools for measuring momentum or for determining overbought or oversold conditions. Because momentum is very different from either a charting technique or a moving average, it is valuable either on its own or as a confirmation of another method.

A word of caution: Trading against the trend can be exciting and profitable, but at considerably greater risk than a trend-following system. The problem with selling an overbought condition is that it is much more difficult to hold losses to a minimum. A long position may be entered while prices are falling fast, and they may continue to fall at the same speed after you are long. Even a quick exit may sustain substantial losses.

%R Method

After the publication of Williams' *How I Made One Million Dollars . . . Last Year . . . Trading Commodities*, the *%R* oscillator became well known. It is a simple way of calculating where today's closing price fits into the recent trading range. Using the last 10 days, define

$$\%R = \frac{Buying\ power}{Range} = \frac{High_{10} - Close_{today}}{High_{10} - Low_{10}}$$

Williams' 10-day *%R* is different from a *10-day stochastic*, because it measures how strong the market closed today compared to the high of the past 10 days. It is also conceptually upside down; that is, as the close gets stronger the value of *%R* gets smaller. It may be intuitively easier to work with if you use 1.0 −*%R*. Williams viewed this as a timing device to add positions within a major technical or fundamental trend. This same approach was discussed with regard to the stochastic, and is shown in Figure 9.16. Trades were not entered if they contradicted the major market direction.

The Ultimate Oscillator

In the *Ultimate Oscillator*, Williams seems to combine his original idea of the A/D Oscillator with a great deal of Wilder's RSI.[12] He adds the unique feature of three concurrent time periods in order to offset the negative qualities of the short time period used for the *%R*, without slowing the system too much. The Ultimate Oscillator uses the following steps:

1. Calculate today's *buying pressure* BP_t by subtracting the *true low* from the closing price, $BP_t = C_t - TL_t$. The true low $TL_t = \min(L_t, C_{t-1})$.
2. Calculate today's *true range*, $TR_t = \max(H_t - L_t, H_t - C_{t-1}, C_{t-1} - L_t)$.
3. Total the buying pressure BP_t separately over the three intervals 7, 14, and 28 days, designated as SB_7, SB_{14}, and SB_{28}.
4. Total the true range TR_t over the same three periods, SR_7, SR_{14}, and SR_{28}.
5. Divide the sum of the buying pressures by the corresponding true range, that is, SB_7/SR_7 and scale by multiplying the 7-day value by 4 and the 14-day value by 2. All three calculations are now in the same scale.

Notice that the nearest seven values for the buying pressure and the true range are each used seven times, that is, they are multiplied by both the scaling factors of 4 and 2, and used once more in the 28-day calculation. Williams has created a step-weighted momentum, assigning values of 7, 3, and 1 to the first 7 days, second 7 days, and last 14 days, respectively. The last 14 days account for only 10% of the total.

The rules for using this oscillator (Figure 9.19) are:

1. A *sell set-up* occurs when the oscillator moves above the 50% line, peaks at a high value, declines, and then moves higher. If the oscillator fails to move above the peak on the next rally, a *short sale* order can be placed when the oscillator fails on the right shoulder. This is a traditional top confirmation signal.

FIGURE 9.19   Williams' Ultimate Oscillator.

---

[12]Larry Williams, "The Ultimate Oscillator," *Technical Analysis of Stocks & Commodities* (August 1985).

2. Short positions are closed out when a long signal occurs, when the 30% level is reached, or if the oscillator rises above 65% (the *stop-loss* point) after being below 50%.

3. A *buy* signal is given using the opposite formation as the short signal (rule 1).

4. Close out longs when a short signal occurs, when the 70% level is reached, or if the oscillator falls below 30% (after being above 50%).

Relative Vigor Index

John Ehlers, who has contributed extensively in the mathematical analysis of prices, in particular using cycles, has created the *Relative Vigor Index* (RVI), a very smoothed momentum indicator.[13] The basic form of RVI is

$$RVI = (Close - Open)/(High - Low)$$

However, the final RVI uses a 4-day symmetric weighting (similar to a triangular weighting) of the *close – open* in the numerator, and a similar symmetric weighting of the *high – low* in the denominator.

   RVI is conceptually similar to the A/D Oscillator; however, the RVI is smoothed in a special way that targets a particular price cycle and eliminates the 2-bar cycle and its associated unwanted frequencies. While it is preferable that the price series be analyzed for its cycle period, Ehlers suggests using 10 as the nominal value. The RVI is calculated as

$$N_t = [(C_t - O_t) + 2 \times (C_{t-1} - O_{t-1}) + 2 \times (C_{t-2} - O_{t-2}) + (C_{t-3} - O_{t-3})]/6$$

$$D_t = [(H_t - L_t) + 2 \times (H_{t-1} - L_{t-1}) + 2 \times (H_{t-2} - L_{t-2}) + (H_{t-3} - L_{t-3})]/6$$

$$Numerator_t = \Sigma\, N_i, \; i = t - n + 1, \, t$$

$$Denominator_t = \Sigma\, D_i, \; i = t - n + 1, \, t$$

$$RVI = Numerator_t/Denominator_t, \text{ while } Denominator_t \neq 0$$

$$RVI\ signal\ line_t = (RVI + 2 \times RVI_{t-1} + 2 \times RVI_{t-2} + RVI_{t-3})/6$$

where O_t, H_t, L_t, and C_t are today's open, high, low, and closing prices and n is the calculation period, nominally 10.

   The RVI signal line is used in the same manner as the MACD signal line. After a peak in the RVI value, showing an overbought situation, the sell signal occurs the first time that the RVI crosses the RVI signal line moving lower.

---

[13]John F. Ehlers, "Relative Vigor Index," *Technical Analysis of Stocks & Commodities* (January 2002).

DOUBLE-SMOOTHED MOMENTUM

Important contributions to the study of momentum have been made by William Blau.[14] In addition to creating new momentum indicators, he has added substantial value to the old ones. Also refer to his work on double smoothing of momentum in Chapter 7.

True Strength Index

Much of Blau's work combines double smoothing of momentum values (1-period price differences) which has surprisingly little calculation *lag* given the amount of smoothing. By using the first differences, he has based the calculations on values more sensitive than price and then slowed them down by smoothing. In effect, he speeds up the price movement before slowing it down. The net result is that the final index value has less lag than we would normally expect, and the index line is much smoother than a standard moving average. Blau refers to this as *using momentum as a proxy for price.* One of Blau's most popular indicators is the *True Strength Index* (TSI) which combines these features:

$$TSI\,(close, r, s) = \frac{100 \times XAverage(XAverage(close - close[1], r)s)}{XAverage(XAverage(absvalue(close - close[1], r)s)}$$

Where r = the calculation period of the first momentum smoothing
         s = the calculation period of the second momentum smoothing
         $close - close[1]$ = the 1-day momentum
         $XAverage(close, period)$ = the TradeStation function for exponential smoothing
         AbsValue = the TradeStation function for absolute value

A spreadsheet to calculate the True Strength Index for crude oil, *TSM True Strength Index*, and a program indicator, *TMS True Strength*, are available on the Companion Website. The spreadsheet code, which is easily entered, is set up as follows:

1. Column *B* is the closing prices
2. Column *C* is the 1-day difference in prices
3. Column *D* is the first smoothing, where D3 = C3 and D4 = D3 + $H$2*(C3-D3)
4. Column *E* is the second smoothing, E4 = D4, and E5 = E4 + $H$3*(D5-E4)

The smoothing constants in *H* are derived from the days in *G* using $2/(n+1)$. A sample of the spreadsheet appears in Table 9.3.

The numerator and denominator of the TSI differ only in that the denominator takes the absolute value of the price changes (the 1-day momentum). This guarantees

---

[14] William Blau, *Momentum, Direction and Divergence* (New York: John Wiley & Sons, 1995).

TABLE 9.3 TSI Calculations Using Two 20-Day Smoothing Periods

	A	B	C	D	E	F	G	H
Column	Date	Crude	Diff	Smooth 1	Smooth 2			
2	1/3/2011	96.35				Smooth 1	20	0.0952
3	1/4/2011	94.18	−2.17	−2.170		Smooth 2	20	0.0952
4	1/5/2011	95.1	0.92	−1.876	−1.876			
5	1/6/2011	93.18	−1.92	−1.880	−1.876			
6	1/7/2011	92.83	−0.35	−1.734	−1.863			
7	1/10/2011	94.05	1.22	−1.453	−1.824			
8	1/11/2011	95.91	1.86	−1.137	−1.758			
9	1/12/2011	96.42	0.51	−0.980	−1.684			
10	1/13/2011	95.85	−0.57	−0.941	−1.613			

that the denominator will be at least as large as the numerator. The 1-day differences are first smoothed over the period r, and then the result is smoothed over the period s. The relationship between the standard momentum (the difference in prices over r days) and the TSI can be seen in Figure 9.20 for Intel (INTC). The standard 20-day momentum indicator (second panel) has the typical erratic pattern of prices, and a slight lead identifying the peaks. The TSI is much smoother with peaks and valleys lagging prices slightly (third panel). If more smoothing is necessary to avoid false signals, a *signal line* can be created by smoothing the TSI using a 3-period moving average, then buying when the TSI crosses the signal line after a high or low value was reached. This can be a basic trend signal, or it can be applied in the manner of MACD. The slight lag in the TSI seems a small problem compared to the extreme noise of the momentum calculation.

Additional Smoothing without Adding Lag

In creating the TSI, Blau missed an opportunity to improve the smoothing with only a minor increase in the lag. Instead of taking the 1-day differences, substitute the n-day differences in the first step. This smoothes the trendline even more at the cost of a slight additional lag. Figure 9.20 shows the TSI with a 10-day difference followed by two 20-day exponential smoothings in the bottom panel.

Anticipating the Turn

When working with trendlines that are very smooth, such as the 10-20-20 TSI, you can anticipate the change in the trend direction most of the time. Instead of waiting for the smoothed trendline to change from up to down, you can sell when it gets to a "near-zero slope" and is continuing to flatten. This anticipation can greatly reduce the lag and improve performance even at the cost of a few false signals.
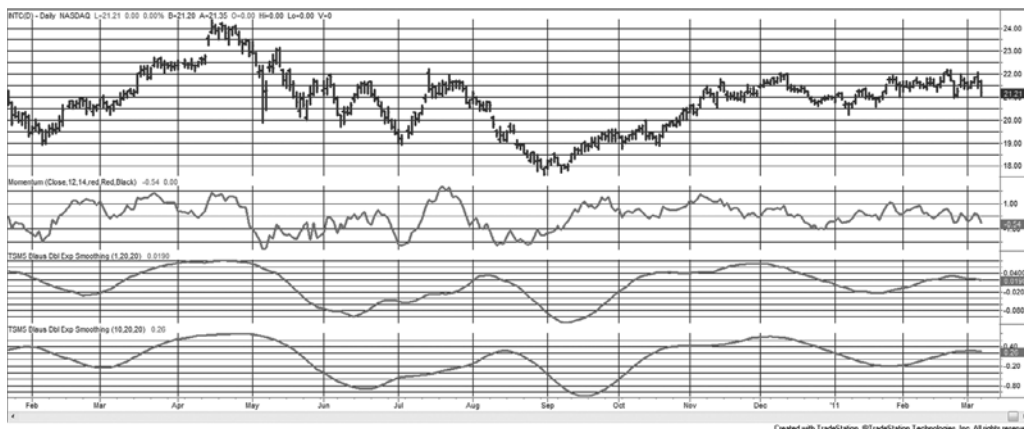
FIGURE 9.20 Comparing the TSI with 10-20-20 smoothing (bottom) to a standard 20-period momentum (center) using INTC from January 2001 through March 2002.

Double-Smoothed Stochastics

Because of Blau's great interest in double smoothing, he defines the general form of a *double-smoothed stochastic* as:

$$DS\,(q,r,s) = \frac{100 \times XAverage\,(XAverage\,(close - Lowest(low,q),r),s)}{XAverage(XAverage(Highest(high,q) - Lowest(low,q),r),s)}$$

Where   *Close – Lowest(low,q)* = the numerator of Lane's raw stochastic, the lowest low over the past *q* periods

     Highest(high,q) – Lowest(low,q) = the denominator of Lane's stochastic, the greatest high-low range over the past *q* periods

     XAverage((...,r),s) = an exponential smoothing of the numerator, first calculated over *r* periods, then over *s* periods

TRIX

Similar to Blau's double smoothing is *TRIX*, a triple-smoothed exponential that is most often used as an oscillator. Introduced by Jack Hutson,[15] it is created using steps similar to Blau except that there are three exponential smoothings and the differencing is done at the end. Typically, the same smoothing constants (calculation periods) are used for each smoothing. This method has been applied to daily, hourly, or even 1-minute price data.

---

[15] Referenced in Robert W. Colby, *The Encyclopedia of Technical Market Indicators* (New York: McGraw-Hill, 2003) as "Good Trix" by Jack K. Hutson, *Technical Analysis of Stocks & Commodities* 1, no. 5.

1. Calculate the natural log (ln) of the closing prices (daily or intraday bars). This implicitly corrects for price volatility; however, it is commonly omitted from the calculation because back-adjusted data in futures will cause errors.

2. Calculate the p-period exponential smoothing of the closing prices, or the ln of the closing prices, to get *trend* 1.

3. Calculate the q-period exponential smoothing of *trend* 1 to get *trend* 2.

4. Calculate the r-period exponential smoothing of *trend* 2 to get *trend* 3.

5. Get the 1-period differences of *trend* 3 by subtracting each value from the previous value. As with the added smoothing of the TSI, the 1-period differences can be replaced with the *s*-period differences.

6. Scale the results by multiplying by 10,000. This is an attempt to get TRIX scaled to a positive integer value for charting and may also be omitted.

The resulting TRIX indicator acts as an oscillator due to Step 5. It can be very smooth when the calculation periods are larger; it also reduces the lag because of the differencing step. TRIX can be used as a trend indicator by buying when the value of TRIX crosses above zero and selling when it crosses below zero. It can produce buy and sell signals sooner by buying when the TRIX value is rising for two or three consecutive periods, and selling when TRIX is falling for two or three consecutive periods. Because the triple smoothing results in a very smooth TRIX value, trading signals can safely use the change in TRIX as an advance indicator of trend.

Figure 9.21 compares TRIX (center panel) with the TSI (lower panel), both based on first differences, but using the smoothing constant of 1.0 to negate one of the steps. The point is to show that TRIX is actually smoother because the differences are taken at the end rather than the beginning. The lag in TRIX is slightly more than the TSI, but it would be unnecessary to use n-day differences for TRIX to get further smoothing; therefore, TRIX may have less lag in its final form.
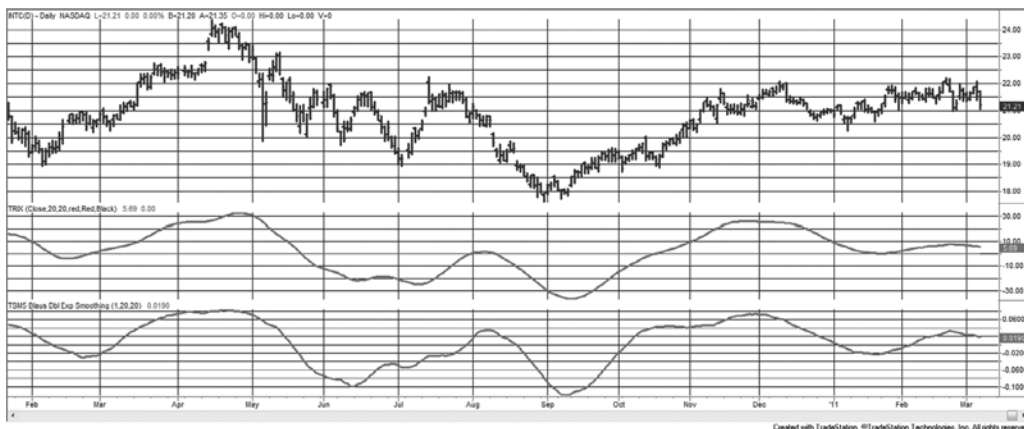
FIGURE 9.21 Comparison of TRIX (center panel) and TSI (lower panel) using two 20-day exponential smoothings applied to INTC.

Changing the Divisor

All return calculations divide the price change from yesterday to today by the starting value, yesterday's price. There is no rule that says that you cannot divide the change by the current price,

$$r_t = \frac{Close_t - Close_{t-1}}{Close_t} - 1$$

While this would result in only small changes, the cumulative effect is said to add stability to an indicator that is based on returns.

An Oscillator to Distinguish between Trending and Sideways Markets

The lack of predictability of trending markets is the greatest problem for the analyst. The work found under the topics "Ranking of Markets for Selection" and "Directional Movement" in Chapter 23 discusses that issue. Based on the idea that the trend component is stronger when price is further from fair value, and the noise (sideways movement) is greater when price is near value, an oscillator can be created to show the strength of the trend component based on this concept,

$$Strength\ Oscillator_t = \frac{Average(Close_t - Close_{t-1}, n)}{Average(High_t - Low_t, n)}$$

As the trend increases, the average change in closing prices becomes larger relative to the high-low range for the day. During an unusual period, when the market gaps open, it would be possible for the differences in the closing prices to become larger than the daily range. In a sideways market, both the change in the closes and daily range will get smaller, but the net price change over period n should be close to zero. This oscillator can be smoothed by taking the change in price over two or three days (for example, $close_t - close_{t-3}$), rather the most recent day, as well as taking the high-low range over the same number of days. The indicator and function, *TSM Strength Oscillator*, can be found on the Companion Website.

Adding Volume to Momentum

A momentum indicator can also incorporate volume by multiplying the price change over n periods by the total volume over that same period. The use of the cumulative volume over the period, or even better, the average volume, will help to stabilize the volume, which is often erratic when seen as only one day's activity. The *average volume* will appear to be the same magnitude as the volume and can be plotted along with it on a

chart. That gives the *momentum-volume indicator (MV)*, shown mathematically and in programming notation:

$$MV_t = (p_t - p_{t-n}) \times \frac{\sum_{i=t-n+1}^{t} volume_i}{n}$$

```
MV = (close-close[n])*average(volume,n)
```

Alternately, the price change over *n* periods could have been divided by *n* to give a per unit value. The following sections will include those techniques that combine price change and volume or open interest; for methods that use volume but not price, see Chapter 12.

Scaling by a Percentage or Volatility

The same conversions can be applied to momentum with and without volume. Using a percentage rather than price will add some robustness over long test periods. Because volatility often increases faster than a fixed percentage of the price when prices rise, momentum can be scaled according to a shorter measure of true range. If the true range is averaged over 20 to 65 days, approximately one month to one quarter, then the 1-day change in price will become a relative momentum value. By using a much longer period for averaging the true range, you can create a stable profile of the volatility changes in the underlying market.

Percentage momentum with volume
```
%MV = (close − close[n]) / close[n] * average(volume,n)
```
Momentum with volume scaled by true range
```
TRMV = (close − close[n]) / average(truerange,p) * average(volume,n)
```
where *truerange* is always calculated for the most recent period (e.g., 1 day), and the average of the 1-day true ranges for the past *p* days is *average(truerange,p)*.

Volume-Weighted RSI

In the same way that the RSI accumulates the price changes for positive days and divides by the sum of the negative changes, it is possible to weight each day by its volume to add another factor, called *money flow*,[16] to the calculation. A positive or upwards day is when the average of today's high, low, and close is greater than the previous average. Each average is then multiplied by the volume, giving the daily money flow, and a ratio of the past 14 days is used to create a *money ratio* and finally a *money flow index*, both steps similar to Wilder's RSI.

---

[16]Gene Quong and Avrum Soudack, "Volume-Weighted RSI: Money Flow," *Technical Analysis of Stocks & Commodities* (March 1989).

$$Moneyflow_t = Volume_t \times \frac{High_t + Low_t + Close_t}{3}$$

$$Moneyratio_t = \frac{\sum_{i=t-13}^{t} Moneyflow_i \, if > 0}{\sum_{i=t-13}^{t} Moneyflow_i \, if < 0}$$

$$Moneyflowindex_t = 100 - \frac{100}{1 + Moneyratio_t}$$

Herrick Payoff Index

Using the change in the underlying value of the futures contract, rather than only the change in price, the *Herrick Payoff Index*[17](HPI) combines volume and open interest to generate an indicator that is not bounded as is the basic momentum calculation. The daily value is:

$$HP_t = cf \times V_t \times (M_t - M_{t-1}) \times \left[1 + \left(\frac{M_t - M_{t-1}}{| M_t - M_{t-1} |} \right) \left(\frac{2 \times | OI_t - OI_{t-1} |}{\min(OI_t, OI_{t-1})} \right) \right]$$

Where

t = today
$t - 1$ = the previous day
cf = the conversion factor (value of a one big point move)
V_t = today's volume
$(M_t - M_{t-1})$ = the difference in the mean values, $M = (high + low)/2$, where vertical bars denote absolute value
$|OI_t - OI_{t-1}|$ = the absolute value of the change in open interest (for futures)
$\min(OI_t, OI_{t-1})$ = the smaller of the open interest for today or the previous day

The expression that divides the change in mean prices by the absolute value of the same change is used to create a value of +1 or −1. The index HP_t is then scaled down to a manageable value and smoothed using a 0.10 smoothing factor, s (about 19 days). This complex formula for HPI can also be written in programming code as

```
HP = BigPointValue*volume*((high-low)/2 – (high[1]-low[1])/2)*
        (1+(((high-low)/2 – (high[1]-low[1])/2) /
      absvalue((high-low)/2 – (high[1]-low[1])/2))*2*
        (absvalue(opint – opint[1])/lowest(opint,2)))

HPI = smoothedaverage(HP,19)
```

---

[17] From the original *CompuTrac* manual, which became the Dow Jones Telerate division, and finally just Dow Jones.
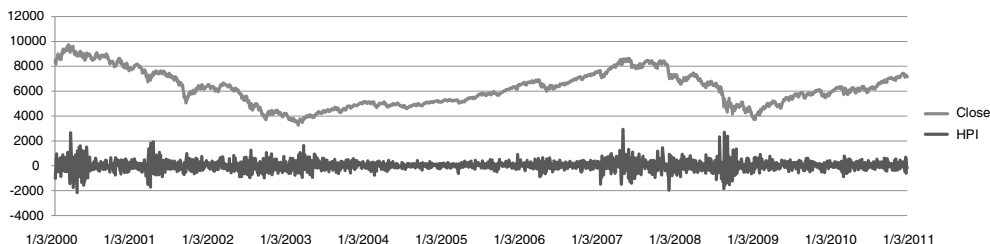
FIGURE 9.22 The Herrick Payoff Index applied to the DAX, 2000 through February 2011.

Most analysts who use the Herrick Payoff Index divide the HPI by 100,000 to scale the value to a more usable level (in the example below it is scaled by 1,000,000). The final series, when seen along with prices, may appear volatile and require interpretation, often using trendlines. This is due to the fluctuations in volume and open interest, which are smoothed over 20 days, rather than a longer period. The Herrick Payoff Index may be helpful, despite its volatility, because it is a combination of factors not included in most other indexes. It has patterns that appear to lead price change to compensate for its noisy behavior. A spreadsheet, *TSM Herrick Payoff Index for DAX*, and a program, *TSM Herrick Payoff*, can be found on the Companion Website.

Figure 9.22 gives an idea of how the HPI reacts to prices, in this case using the German DAX from 2000 through February 2011. It shows higher volatility at the turning points and very low volatility during the upwards trends from 2004 to 2006 and again in 2010. It turns out that volume is a good surrogate for volatility and might actually be a good forecaster of volatility. This is discussed in Chapter 12.

Comments on the Use of Volume

Volume is an important piece of information, but it can be difficult to interpret. It fluctuates in a much larger range than price, and may be 50% higher or lower from day to day. While it indicates market interest and potential volatility, there are many days for which a high or low volume does not have a consistent reaction.

In general, adding volume to an indicator results in a more volatile, erratic series. Therefore, the first steps in using volume are:

1. Create a long-term, smoothed volume series.
2. Locate days with extremely high volatility to identify only those exceptional days where continued high volatility should follow.
3. Low volume should not be determined by a single day, but by either a few unusually low days clustered together or by a decay in the smoothed volume over a modest time period.

Using volume to enhance trend signals as well as identifying important volume patterns is discussed in Chapter 12.

VELOCITY AND ACCELERATION

In physics, momentum is called *speed* or *velocity*. It is the amount of change measured over a specific period of time. It is also called *rate of change* and with regression analysis it is *slope*, discussed in detail in Chapter 6.

$$Slope_t = \frac{p_t - p_{t-n}}{x_t - x_{t-n}}$$

where the numerator is the change in price and the denominator is the change in the *x*-axis, which is normally the time, or simply n, elapsed units. For example, when traveling in a car, your speed might be 30 miles per hour. When prices are moving higher, they may be averaging a gain of $1 per week.

There are two types of velocity: *average* and *instantaneous*. The *average velocity* is simply calculated as the mean velocity over a fixed distance and for a fixed time interval. In working with stock and futures prices, the time interval used is days and the distance is measured in points or unit value; if IBM moved $5 in six days, its average velocity is

$$\overline{v} = \frac{5}{6} = \$0.833 \ \ per \ \ day$$

In general, the average velocity is expressed

$$\overline{v} = \frac{D}{T}$$

where D = the total elapsed distance over the time interval T

Velocity is the same as the simple measurement of momentum. For a geometric interpretation of momentum, the change in price, D, can be related to the length of the momentum span, T, giving the same results for average velocity as for slope. Physicists prefer to draw velocity (the average over period T) and instantaneous velocity (the speed at exactly one point) in their own way, shown in Figure 9.23.

The *instantaneous velocity*, v, which is the velocity calculated at a specific point in time, may be different from velocity. In order to determine the instantaneous velocity, a mathematical technique called *differentiation* is used. It effectively looks at smaller and smaller time intervals and consequently smaller distances on the price curve until the slope calculation is reduced to a single point. The result of the process of differentiation is called the *derivative* and is expressed

$$v_t = \lim_{\Delta_t \to 0} \frac{\Delta D}{\Delta t} = \frac{dD}{dt}$$

This shows that the velocity taken at any point is the result of the time interval (t) becoming progressively smaller without actually reaching zero. The symbol delta (Δ) represents

(a)

(b)
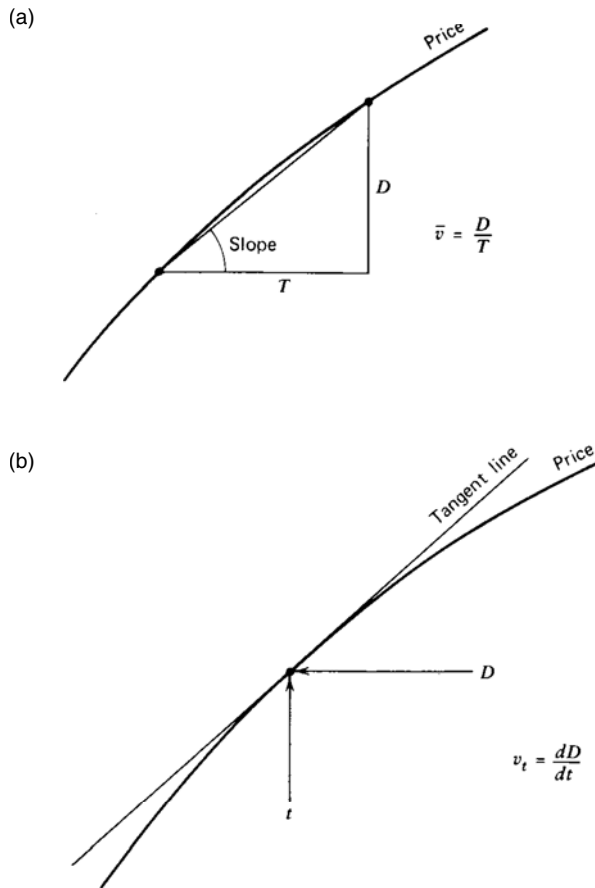
FIGURE 9.23   (a) Average velocity. (b) Instantaneous velocity.

the change in price (ΔD) and the corresponding change in time (Δt). The rules for differentiation can be found in any advanced mathematics book. Only the results are presented here. The velocity v_t represents the speed or momentum of the price at the point in time t. If v gets larger for t_0, t_1, t_2, . . ., then the velocity is increasing; if v gets smaller, the velocity is decreasing. Because the velocity also denotes direction, it can be both positive and negative in value and appear similar to a momentum indicator.

   Acceleration is the change in velocity. In the same way that we find the change in price D over time period t, we can find the change in velocity over time period t. Therefore, if you are driving at 30 miles per hour when you enter the acceleration ramp of the motorway, and after one minute you are driving 60 miles per hour, you have changed speed, or accelerated, at the rate of 30 miles per hour per minute. Fortunately, we do not need to be concerned about the units of time when we apply these techniques to prices. The units are always the same, whatever they are. When S&P prices have been moving

higher at an average speed of 10 points per week and then begin posting increases of 15 points per week, then 20 points per week, prices are accelerating at the rate of 5 points per week. Mathematically, velocity can be substituted for price in the equation for v_t to get acceleration, a_t.

$$a_t = \lim_{\Delta t \to 0} \frac{\Delta V}{\Delta t} = \frac{dV}{dt}$$

The short-cut to finding velocity and acceleration is to take the first and second differences, each step removing first the trend, then the speed, leaving the more sensitive components plus noise.

Finding the Velocity and Acceleration of Different Techniques

Differentiation can be applied to many different formulas that have been discussed previously, including those that represent a straight line, curved lines, and various other trendlines. The result of the first differentiation gives you the component that represents velocity, and the result of the second differentiation is the component of acceleration. Of course, some of the basic equations have constant velocity and cannot be used for a velocity trading plan because the values never change. The straight line, simple and weighted moving averages, and exponential smoothing all have constant velocities. Only those equations with *second-order smoothing* will work. Table 9.4 gives basic equations along with their first and second derivatives.

TABLE 9.4 Equations for Velocity and Acceleration

	Basic Equation	Velocity at X_t^*	Acceleration at X_t^*
Straight line	$y_t = a + bx_t$	$v_t = b$	$a_t = 0$
Curvilinear	$y_t = a + bx_t + cx_t^2$	$v_t = b + 2cx_t$	$a_t = 2c$
Logarithmic (base a)	$y_t = \log_a x_t$	$v_t = (\log_a e)/x_t$ or $v_t = 1/(x_t \ln a)$	$a_t = -1/(x_t^2 \ln a)$
Logarithmic (natural log)	$y_t = \ln x_t$	$v_t = 1/x_t$	$a_t = -1/x_t^2$
Exponential	$y_t = e^{ax_t}$	$v_t = ae^{ax_t}$	$a_t = a^2 e^{ax_t}$
Moving average	$y_t = \dfrac{x_t + x_{t-1} + \cdots + x_{t-n+1}}{n}$	$v_t = 1$	$a_t = 0$
Weighted moving average	$y_t = \dfrac{a_1 x_t + a_2 x_{t-1} + \cdots + a_n x_{t-n+1}}{n}$	$v_t = \dfrac{a_1 + a_2 + \cdots + a_n}{n}$	$a_t = 0$
Exponential smoothing	$y_t = y_{t-1} + c(x_t - y_{t-1})$	$v_t = c$	$a_t = 0$

*Because velocity and acceleration are time derivatives, all equations implicitly include the factor $\dfrac{d(x_t)}{d_t}$ as part of the right member.

Let's assume that the velocity and acceleration have been calculated as in Table 9.4. The following are the possible combinations that can occur:

Velocity	Acceleration	Price Movement
+	+	Price is moving up at an increasing rate
+	0	Price is moving up at a constant rate
+	–	Price is moving up at a decreasing rate
0	0	Price is static
–	+	Price is moving down at a decreasing rate
–	0	Price is moving down at a constant rate
–	–	Price is moving down at an increasing rate

Using acceleration, a change of velocity (or momentum) can be detected or the strength of a current price move can be confirmed.

Using Velocity and Acceleration to Identify a Sideways Market

The combination of velocity and acceleration can give a good indication of whether prices are moving in a directional or sideways pattern. When velocity is near zero, the average speed of price movement is near zero. It is also the same as saying that prices have netted no movement over the time interval T. However, that is not enough.

During the interval T, prices may have moved sharply higher then sharply lower. By chance, they are unchanged after time T although they are still moving lower quickly. The acceleration will tell you that prices are moving even though they are at the same level as T periods ago. To identify a sideways pattern, both the velocity and acceleration must be near zero.

Quick Calculation of Velocity and Acceleration

A less precise but very convenient way of isolating velocity and acceleration is the calculation of *first* and *second differences.* The purpose of these values is to find more sensitive indicators of price change, and most traders find this quick calculation satisfactory. The results can be used in exactly the same way as the formal mathematical results. Consider the following two examples:

1. A price series 10, 20, 30, 40, . . . is moving higher by a constant value each day. The first differences are 10, 10, 10, . . ., showing a consistent velocity of 10. The second differences, formed by subtracting sequential values in the first-difference series, are 0, 0, 0, . . ., showing that there is no change in speed; therefore the acceleration is zero.

2. Another price series is shown with its first and second differences as

Time	1	2	3	4	5	6	7	8	9	10	11	12
Series	10	15	20	30	45	50	45	35	25	20	25	40
Velocity		+5	+5	+10	+15	+5	–5	–10	–10	–5	+5	+15
Acceleration			0	+5	+5	–10	–10	–5	0	+5	+10	+10

where velocity values are the first differences, and acceleration the second differences. The original series has two turns in the trend direction clearly shown by the velocity and acceleration as changes in the sign of the numbers. The velocity continues to be positive through the sixth value as the underlying price moves from 10 to 50. Whenever prices change direction, the velocity changes sign. The basic upward trend can be associated with a positive velocity and a downward trend with a negative one.

The acceleration shows the change in speed. At the sixth item, the acceleration becomes negative, even though the velocity was positive, because prices moved higher at a slower rate. They had been gaining by 5, 10, and 15 points each day, but on day 6 the gain was only 5 points. This reversal in acceleration was a leading indicator of a trend change. A similar situation occurred on day 8, when the acceleration slowed and reversed on day 10, one day ahead of the actual price reversal.

HYBRID MOMENTUM TECHNIQUES

Combining a Trend and an Oscillator

Directional Parabolic System

The *Directional Parabolic System*[18] is a combination of two of Wilder's well-known techniques, *Directional Movement* and the *Parabolic Time/Price System*. Directional Movement is covered fully in Chapter 23. It gained popularity as a method of selecting the futures markets that were most likely candidates for trend-following systems. The Parabolic Time/Price System is covered in Chapter 17. Although a full reading of both techniques is necessary, the essence of the combined systems may be understood with the following definitions:

+DM14	The 14-day upward Directional Movement, the sum of the current high prices minus the previous high prices.
−DM14	The 14-day downward Directional Movement, the sum of the previous low prices minus the current low prices.
ADX	The Average Directional Movement Index, calculated by smoothing the Ratio of the net of +DM14 and −DM14 by the sum of the same values.
DPS	The Directional Parabolic stop-loss.

Although shown as −DM14, the downward Directional Movement is a positive number based on the sum of those days that closed lower. The ADX is a ratio of the +DM14 divided by −DM14 representing the positive direction of the index. Therefore, the ADX is an oscillator such that, when its value is greater than 50, it means that price movement

---

[18]J. Welles Wilder, Jr., *Chart Trading Workshop 1980* (Greensboro, NC: Trend Research, 1980).

is upwards. Directional Movement is combined with the Parabolic Time/Price System according to the following rules:

1. If the ADX is up, take only long Parabolic System trades; if the ADX is down, take only short trades.
2. If the two systems conflict, no trade is entered. A trade may be entered at the time they agree.
3. If a position is held, the Parabolic stop is used. (The stop is now called the DPS instead of the SAR because it no longer requires a reversal of position.)
4. If no position is held, the Directional Movement *equilibrium point* is used (the high or low of the day on which the +DM14 crosses the −DM14).

In 1980, the entry rules were revised to include an added use of the ADX when it is greater than the +DM14 or the −DM14. Because the ADX is an oscillator and indicates turning points in the trend, when the ADX exceeds the magnitude of the current +DM14 or −DM14 and reverses, the current position should be closed out. If the ADX remains above both the +DM14 and −DM14, the market is extremely strong, and liquidation should stop. The ADX is intended to be a leading indicator for liquidation only. Reversal of the current position only occurs when the DPS has been penetrated and the new trade agrees with the direction of the Parabolic System.

The addition of an oscillator to a trend-following system allows trades to be closed out at more favorable prices than the usual trend exits. If the new direction carries through and the position is reversed, the added feature has worked perfectly; however, if prices turn back in the original direction, a reentry may not be possible. The revised rules are unclear concerning reentry into a position if prices fail to penetrate the DPS and signal a reversal. A reentry might occur if the ADX falls below both the +DM14 and −DM14, indicating that prices are no longer extreme, then turns back in the trend direction. Once reestablished, the DPS can be used and additional exits using the revised rules would apply.

Oscillator Method with ADX Filter

A simple variation from Lars Kestner[19] creates an oscillator from two moving averages, then uses the ADX to show a lack of trend. The rules are

Osc = average(close,10) – average(close,50)
If osc < osc[1] and ADX(14) < 30 then sell next open
If osc > osc[1] and ADX(14) < 30 then buy next open
If osc > osc[1] then buy to cover next open
If osc < osc[1] then sell next open

---

[19] Lars Kestner, *Quantitative Trading Strategies* (New York: McGraw-Hill, 2003).

Using a Oscillator as Timing for a Trend System

One point of irritation is that trend entries always seem to be poorly timed. Traders would prefer to wait for pullback after a buy or sell signal to enter at a better price. This can be done using an oscillator of a much shorter calculation period than the trend. One choice would be

1. Calculate a 40-day moving average
2. Calculate a 5-day stochastic indicator
3. *Buy* when the 40-day trend is up and the 5-day stochastic < 20
4. *Sell* when the 40-day trend is down and the 5-day stochastic > 80

It may also be better to enter trades using this timing method but exit immediately when the trend changes. Waiting for just the right point to exit may result in very high risk if the market is moving the wrong way. In order to avoid missing a trade entirely, a rule can be added that forces entry after n days following the initial signal.

Cambridge Hook

An indicator that combines Wilder's RSI with other basic indicators is the *Cambridge Hook*.[20] It is intended to identify an early reversal of the current trend by combining three indicators. The following conditions must occur in an existing upwards trend:

- An outside reversal day (a higher high followed by a lower close).
- Wilder's RSI must exceed 60% (moderately overbought).
- Volume and open interest must be increasing.

The result is a high likelihood of a downward trend reversal (the opposite applies to upward trend reversals). Protective stops are placed above the high of the hook on the day that signaled a downward reversal. The program function *TSMCambridgeHook*, available on the Companion Website, returns a "1" when an upwards hook occurs and a "–1" when a downwards hook occurs. These events should be filtered with a trend direction.

Note that using the open interest may be a problem. While volume is available through the day for most futures markets, open interest is not as easily available. Because of that, the program *TSM Cambridge Hook*, on the Companion Website, has the option of not using open interest.

MOMENTUM DIVERGENCE

Divergence occurs when two price series move apart. The Dow Jones Industrials and Dow Jones Utilities are diverging if the Industrials are rising while the Utilities are falling. This divergence between these two markets has always been considered a leading indicator

---

[20] Elias Crim, "Are You Watching the 'Right' Signals?" *Futures* (June 1985).
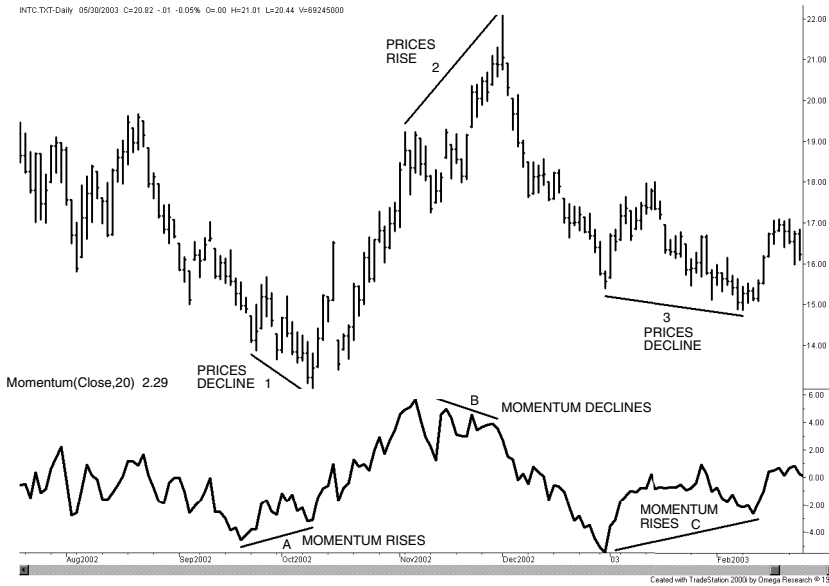
FIGURE 9.24 Momentum divergence.

of a downturn in the economy. The S&P 500 is also watched in relationship to the 10-year Treasury note, the benchmark long-term rate. Notes usually counterbalance moves in the stock market. When the S&P rallies at a fast pace, the price of notes falls to reflect the anticipation of an increase in interest rates needed to dampen growth. When the S&P and the price of notes both rise, or both fall, something special is happening. When prices diverge with respect to a technical indicator, such as an unsmoothed momentum or the MACD, *the direction of prices is expected to follow the direction of momentum.* You can visualize the price chart needed for divergence as the rising part of a rounded top. Prices are still going up but at a slower and slower rate.

Momentum divergence is measured by comparing the direction of prices with the direction of a momentum indicator over the same time interval. Most often, this is done by connecting the peaks of the price movement when prices are rising (or the valleys of the price declines). Connecting only the peaks and valleys of both prices and momentum avoids the problems associated with erratic data. Figure 9.24 shows three examples of a 20-day momentum divergence for Intel.

A *bearish divergence* is one that anticipates a downturn in prices. Because momentum is a leading indicator, a bearish divergence occurs when prices are rising and the momentum values are falling. This can be seen in the middle of Figure 9.24 where line 2 (top panel) shows sharply rising prices at the same time that line *B* (bottom panel) shows clearly falling momentum values. Price follows momentum, and a sharp decline lasts for all of December 2002.

A *bullish divergence* is formed when prices are declining while momentum values are rising. Two examples of a bullish divergence can be seen in Figure 9.24 marked with *1* and *3* on the price chart and *A* and *C* on the momentum indicator.

The important points to remember about a divergence are:

- Prices and momentum must be moving in opposite directions. It is not correct to say there is a bullish divergence when momentum is rising quickly while prices are moving sideways or slightly higher. They *must* be moving in opposite directions.
- The greater the divergence, the more likely prices will change direction soon. That is, when prices are moving up very sharply and momentum is clearly moving lower, then the likelihood of an immediate change of direction for prices is much greater than if prices were gradually rising while the momentum was slowly falling.
- A bearish divergence can be interpreted as a market that is rising slower and slower. Each successive peak makes a smaller advance, and each successive peak occurs after more and more time has elapsed. This may appear similar to a rounded top before prices start down.
- Divergence that occurs over a longer time period (for example, months) will forecast a larger price reversal than a divergence formed over hours or days.

Divergence is best when the momentum indicator begins at an extreme high or low level, indicating that the price move is extended. It is particularly good if the divergence signal occurs while momentum is still well above or below the midpoint of the indicator values, either 50 or 0. That assures us that a price correction can occur before the indicator returns to neutral.

An Amazon.com Example Using Momentum Peaks

Momentum divergence is an important concept. The following example, applied to Amazon.com in Figure 9.25, uses the method of momentum peaks with MACD, plus an additional rule. Use the following steps:

1. Find the swing highs on the chart. This can be done simply by looking at the highest peaks. In Figure 9.25 there are two significant peaks, one in January 1999 and the other at the end of April 1999. There is a peak slightly earlier in April; however, that is part of the price move that ends with a rally to 110.
2. Draw a line connecting the January and April peaks.
3. There will be two corresponding peaks in the MACD lines directly below the price chart. Connect the two peaks in the MACD line.
4. The line drawn across the price highs is clearly rising. The line across the MACD peaks is clearly falling; therefore, the pattern indicates a *bearish divergence.* Prices confirm the divergence by dropping from 110 to 70 in less than two weeks, and then below 50 in the same move.
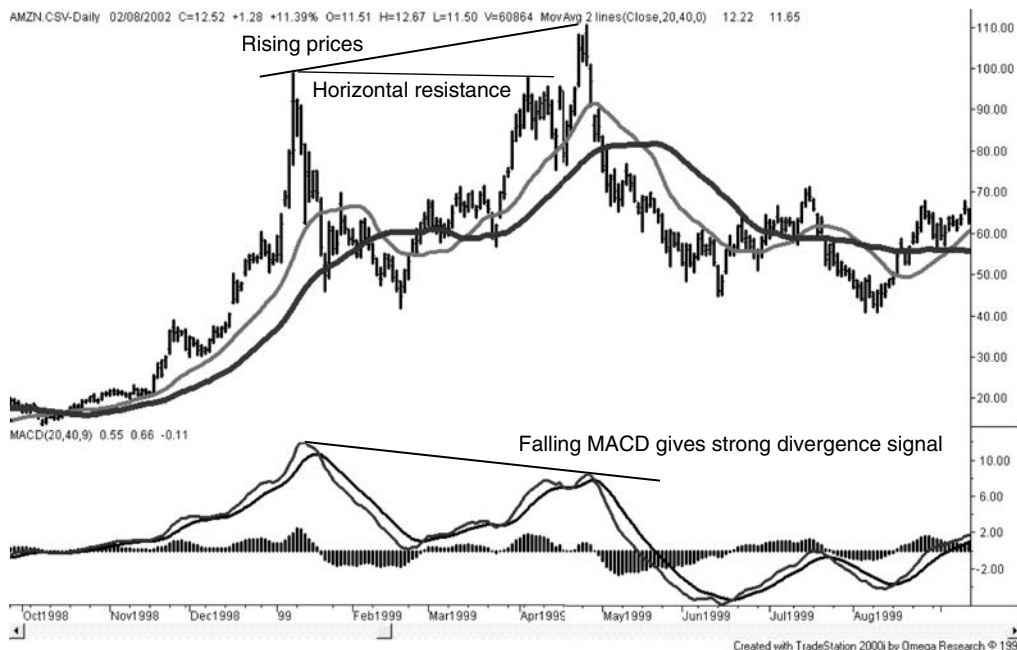
FIGURE 9.25 An example of divergence in Amazon.com.

There is also an unmarked bullish divergence on the chart. Prices bottom in June and August with clearly rising lows in the MACD occurring at the same time. The August price low marks the bottom of the move and a rally follows.

Trading Rules for Divergence

There are a number of alternative rules for trading a momentum divergence, each differing in the amount of anticipation.

MACD Divergence
The simplest rules are based on using the MACD as the indicator to create a bearish divergence. Once the second rising price peak is identified, along with the corresponding MACD peak, the divergence sell signal comes when the MACD line crosses the MACD signal line as it moves lower. This is seen in Figure 9.25 at the end of April. The trade is exited when the MACD value becomes zero, or if a price objective is reached, based on a top formation. It is possible to extend the trade by waiting for the MACD line to cross back over the signal line before exiting. The same rules apply to a bullish divergence.

General Rules for Trading Momentum Divergence
1. *Enter a short position when the divergence is identified, provided prices have not already reached the correction level or profit target.* Bearish divergence is

recognized after the second momentum peak has crested; therefore, it is possible that the momentum value will be near neutral (the midpoint value) if the first momentum peak was not an extreme. The neutral momentum is the normal profit target because we cannot expect high momentum to alternate with low momentum, but we can expect high or low momentum to correct to neutral, or zero. Waiting until the divergence is extremely clear is often too late. Momentum will have achieved most of its correction. An alternative is discussed in the following section, "Anticipating the Divergence."

2. *Enter a short position when the MACD line crosses the signal line after the divergence formation is recognized.* MACD offers a clear signal: the crossing of the faster MACD line with the slower signal line. This is a basic buy or sell signal and applies equally to divergence patterns.

3. *Exit the short position when the current momentum moves above the last momentum peak.* A new momentum high after a divergence signal indicates that the divergence has disappeared and there is no basis for this trade. The exact price at which this occurs may be calculated one day in advance for most momentum indicators.

4. *Exit the short position when the market has corrected or an objective has been reached.* Once the momentum has declined to the midpoint level of 50 for the RSI and stochastic, or zero for the MACD or simple momentum, it should be considered neutral and cannot be expected to continue into negative values. A price objective can also be set using volatility or support levels.

5. *Exit the MACD short divergence when the MACD crosses the signal line moving higher.* MACD provides a signal that may allow the divergence trade to be held longer, or exited quickly. In Figure 9.25 the MACD line gives a sell signal at the beginning of May and does not give another buy signal to close out the trade, until mid-June. This adds considerable profit to the trade.

6. *Allow the short divergence position to convert to a short trend position.* If the MACD is not used, then a simple trend, such as a moving average, can be substituted. A short divergence signal can be converted to a short trend signal using, for example, one of the two trends used to create the MACD.

Anticipating the Divergence

Divergence signals are often seen too late. When the second momentum peak is recognized, especially when the divergence is severe, the momentum values are already near their neutral level, 50 or zero. Anticipating the divergence signal can be a more successful approach to trading.

Bearish divergence can be anticipated at the point where prices move above their previous resistance level. This is shown in Figure 9.25 with the line market *horizontal resistance.* Once prices move higher, there is always a potential divergence. If the current value of momentum is lower than the value of momentum at the previous price peak, an *anticipated divergence* sell signal exists. The short sale is now entered as

prices are rising as long as the current momentum value is below the last peak momentum value. For Amazon.com, that means holding a short position while prices continue higher. The trade is exited if the momentum value continues higher and exceeds the previous peak momentum value, in which case there can be no divergence pattern. This method offers the best opportunity for profiting from the entire downward reversal, but at higher risk. A less risky alternative would be to divide trading capital into three parts, then

- *Sell* the first third when prices make a new high and the MACD value is much lower.
- *Sell* the second part when the MACD value moves to within 15–20% of the previous MACD high.
- *Sell* the third part when the MACD value crosses the signal line heading down.

If there is only one choice, it is better to take the second signal. If there are two choices, take the first and second. If you only take the third sell signal, when the MACD crossed the signal line, prices will have already dropped significantly, and you will be disappointed with your entry price and the lack of profit opportunity.

Single, Double, and Triple Divergences
In fewer cases, double and triple bearish divergences will occur. A double bearish divergence is one in which three momentum peaks are declining with prices rising at each corresponding momentum peak. Most often, the second momentum peak is only slightly lower than the first, and the last peak drops off noticeably, indicating that a drop in price is soon to follow. Multiple divergences are expected to be more reliable than single divergences, and represent a prolonged period in which prices are rising at a slower and slower rate, in the manner of a rounded top.

Alternating Divergence Peaks
A common bearish pattern is where a lower momentum peak falls between two declining peaks. For example, the first momentum peak is at 90, the next at 60, and the last at 75. When studying the price and momentum charts, most analysts will ignore the lower peak in the middle and consider only the 90–75 divergence. In the following section, this combination can be automated by looking at the most recent momentum peak, i, and the previous two momentum peaks, $i - 1$ and $i - 2$, along with their corresponding prices.

Programming Divergence

Before explaining the technique used in the divergence program, it should be clear that there is no simple solution. There are many choices to be made in order to recognize a divergence pattern, and it will be difficult to find them all. What we can see on a chart is not always easy to program into a computer. A TradeStation program, *TSM Divergence*, is available on the Companion Website; it gives the user a number of choices for entries

and exit, including multiple divergences, or simply a starting point for writing your own program. The inputs are

1. *Diverge*   1 = single divergence, 2 = double divergence
2. *Swing*   the percentage price swing to find price extremes (normally 2% to 5%)
3. *Strength*   the minimum percentage decline from the indicator high that will trigger a signal (normally 5%)
4. *Length*   the calculation period for the stochastic (normally 5 to 10)
5. *Type*   0 for normal prices, 1 to change short rates into yields, and 2 to approximate the yield of a bond.
6. *Exit*   the percent added or subtracted from 50 for the momentum exit (*slowK*) criterion (if exit = 10 then exit at 60, normally 0)
7. *FastX*   exit level for fast stochastic (*fastK*) (normally 20)

The first decision is whether the pattern keys off the momentum indicator or the price pattern. This program uses the price peaks, which are identified using the swing technique explained in Chapter 5, with the minimum swing value given in the parameter *swing*. For the momentum indicator, the stochastic is used instead of either a simple momentum or the MACD. Momentum values are then chosen corresponding to those price extremes. The momentum value at the second swing high must be less than the one at the first swing high by the amount specified by the *strength* parameter, typically 5 (for a bearish divergence). There is no test to see if momentum dropped further between the two price peaks. All other momentum values are ignored for the purpose of deciding on the entry signal. If the momentum peaks are declining and the price peaks are rising, there is a bearish divergence. If momentum peaks are rising and the price peaks are falling, there is a bullish divergence. By not looking for the momentum peaks, it is possible that some of the patterns may not be timed correctly.

A bearish divergence exits when the smoothed fast stochastic (%*D*) reaches 50. If the input parameter *exit* is set to +10 or –10 the trade exits when the momentum reaches 60 or 40, respectively. Alternatively, if the momentum drops quickly, the parameter *FastX* will trigger an exit when the raw stochastic, *FastK*, touches 20.

In addition to the standard single divergence, the program recognizes a double divergence, the combination of three rising price peaks and three declining momentum peaks.

It is easier to find divergence by looking at a chart on a quote screen than to program it into a computer. Translating what you see into a systematic analysis of divergence signals is very difficult. You will find that this program does not always find the divergence that seems obvious to the eye. A divergence may be missed when there is a steady rise in prices that do not create swing highs, even though there is a corresponding steady decline in momentum. This situation is addressed using *slope divergence*.

Slope Divergence

One of the problems in using peak prices and peak momentum values is that some of the most obvious divergence situations are missed. Prices can move higher or lower steadily,

without large swings, while momentum moves the other way. This will happen during a very orderly rounded top or rounded bottom formation. Without peaks that can be identified using a swing analysis, this pattern is missed.

An alternative technique is to analyze the slope of both the price movement and the momentum indicator over the same time interval. This can be done using a spreadsheet function, *slope*, or the program function *LinearRegSlope*, over a specified time interval. Because momentum is a way of detrending the price series, the period used for the calculation should not be too long; otherwise the slope values of the momentum will tend towards zero.

A more interesting but complicated method is to record the point of the last momentum peak, m. As we move forward in time to day t, find the slope of the momentum from m to t and the corresponding slope in the price over the same period. If the slope of the momentum is negative and the slope of the prices is positive, then there is a divergence. As t gets larger, we have more confidence in the divergence. If the slope of the prices on day t is less than the slope on day $t-1$, but still positive, we have a sell signal.

Divergence can be any combination of conflicting directions between the slope of price and the slope of momentum, including prices rising faster than momentum, momentum rising faster than prices, or the opposite. However, classic analysis has focused on momentum as a leading indicator of a change in the price direction, which limits the combinations to:

- Prices rising and momentum falling (a *bearish divergence*).
- Prices falling and momentum rising (a *bullish divergence*).

The *strength* of a bearish divergence, which is helpful when selecting which situations are best for trading, can be determined primarily by the momentum slope, but can also be assessed as the net difference between the rising slope of prices and the falling slope of momentum. When comparing the two slopes, care must be taken because the angle of price movement can be far greater than the angle of momentum movement.

Slope Divergence Using Double Smoothing

Double smoothing, discussed earlier in this chapter, is a tool that represents the trend of momentum but may not show many momentum peaks; therefore, it becomes an alternate method for slope divergence. In Figure 9.26 there is a long upwards move in the NASDAQ 100 throughout 1999. The price swings are relatively small and may not be picked up using a swing value that worked during prior years. At the same time there is a steady decline in momentum, represented by a double smoothing of 20-20-20 (a 20-day momentum, smoothed twice using 20-day exponentials). Lines are drawn through both prices and momentum to show the slope of the corresponding movement.

One way to produce a trading signal for the two slope calculations is to monitor their relative movement. While they remain constant or are moving apart, no action is taken. Once the slope values begin to converge beyond a threshold value representing normal variance, a short sale signal is produced. After that, normal price targets apply. If the price slope continues to decline, the trade should be held. If the momentum slope rises
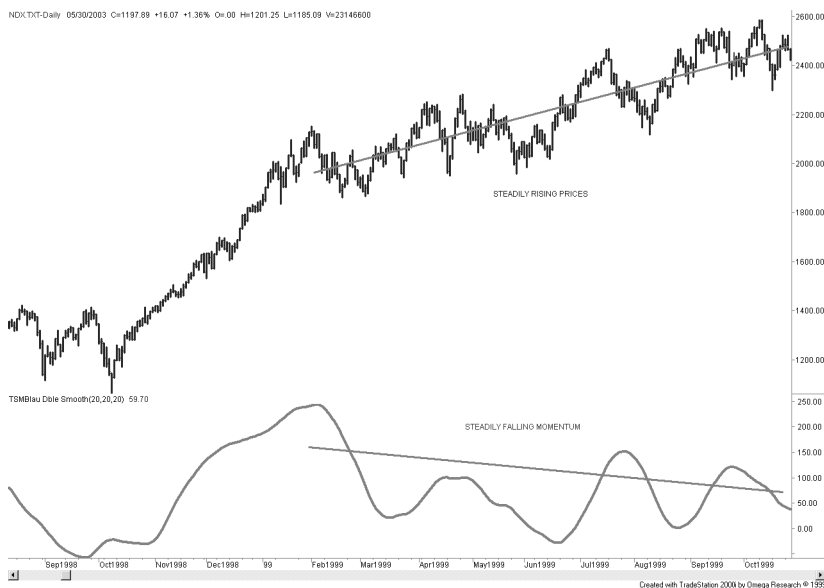
FIGURE 9.26   Slope divergence of NASDAQ 100 using double smoothing.

above its value at the time of the short sale signal, the trade should be exited, or if the slopes begin to diverge significantly, the trade should also be exited.

SOME FINAL COMMENTS ON MOMENTUM

Because momentum and oscillators are very different from either a charting technique or moving averages, they have become important in technical analysis. However, when the time interval for calculation is small, these indicators can be highly unstable, jumping from frequent overbought signals to just as frequent oversold ones. All momentum indicators, when used to *enter* contrary to the direction of price movement, exhibit high risk.

Momentum indicators are most often used as a timing tool within another more conservative strategy, such as trend following. The calculation period is then tuned to allow the momentum values to reach extremes with a certain frequency. Momentum values that are not extremes give us very little useful information. Consider a trending strategy where each trade is held for an average of 20 days. A fast oscillator can be created to provide entry timing. If you are willing to wait up to two days to enter a trade after the trend signal has been given, then construct a 10-period oscillator of 1-hour bars, or a 3-period oscillator of daily bars. Test the oscillator to see if it generates at least one, but preferably two, oversold signals during each 2-day period. If so, use it to time your entry and you are likely to be buying dips rather than rallies.

Seasonality and Calendar Patterns

C hapter 6 introduced prices as a time series and identified its four components as the trend, the seasonal pattern, the cycle, and chance (random) movement; it included various ways of finding the trend using statistical analysis and forecasting techniques. Chapter 7 then showed various ways to calculate trends. Of all techniques, the trend is overwhelmingly the most popular foundation for trading systems. In this and the next chapter, we turn our attention to two other principal components, the seasonal and cyclic movements.

Seasonality is a cycle that occurs yearly. It is most often associated with the planting and harvesting of crops, which directly affect the feeding and marketing of livestock. Normally, prices are higher when a product is not as readily available, or when there is a greater demand relative to the supply, as often occurs with food or heating oil during the winter months and electricity during mid-summer. For grain, cotton, coffee, and other agricultural products, the crop year is dominated by planting, harvest, and weather-related events that occur in between. Most abundant crops have been produced in the northern hemisphere, but South American soybeans and orange juice have become a significant factor since the early 1980s, as have Australian and New Zealand beef and lamb, resulting in a structural change in seasonal patterns. Globalization has not only affected financial markets, but nearly everything we purchase.

Consumer habits can cause a seasonal pattern in metals and some equity markets as weather does for agricultural products. Passenger airline traffic, along with the travel and hotel industry, is much more active in the summer than in the winter, and profits of those companies that are not diversified reflect that seasonality. Gasoline is in high demand during the summer, when most of the population in the northern hemisphere makes room for each other at the beach. Eastman Kodak once had a classic pattern caused by much more active picture-taking during the summer months, which was also reflected in the price of silver, used to make film. Not anymore.

Many commodities are priced in U.S. dollars, such as gold, oil, and even metals traded on the London Metals Exchange (LME), but for investors not in the United States, purchasing in their own currency, the fluctuations in foreign exchange rates significantly change the price of the commodity. Looking at it from the view of a U.S. consumer, if the U.S. dollar falls, then American buyers will expect to pay more. Commodities that have worldwide demand maintain a world price, so that everyone essentially pays the same amount, regardless of their currency. When studying seasonality, it is sometimes necessary to separate the underlying price move of the commodity from the move in its denominated currency.

A CONSISTENT FACTOR

Even when the impact of seasonality on agricultural products is not clear from the price patterns, it is still there. Consider the following factors that do not change:

- More corn, wheat, and soybeans are sold during harvest than at any other time of year because there is not enough available storage to hold all of the new crop. Rental storage, when available, requires a minimum 3-month charge. Lack of storage and the need for immediate income result in greater sales and cause prices to decline.
- Because feed grains are harvested only once each year, forward contracts include a storage cost as part of the total carrying charge. Therefore, each forward delivery price should be higher within the same crop year.

Sometimes the price pattern of forward months does not seem to reflect the added costs of carry. Occasionally these markets even invert, and the nearest delivery trades at a price higher than the deferred months, a situation familiar in crude oil and copper. The cost of carry, however, still exists in an inverted or *backwardation* market. Extreme short-term demand pushes the nearest delivery much higher, while the events causing price disruption are expected to be temporary. The normal carry is still there; it is just overwhelmed by temporary demand.

Seasonal equity shares reflect the same factors as agricultural products. While holiday travel may vary by 10% in a given year, there is still a strong seasonal pattern in the travel business. The profitability of a company may decline during a poor travel year, sending share prices lower, yet the seasonality is still there.

It is important to be able to identify seasonal patterns. Seasonal patterns can bias the size of the positions traded throughout the year, they can identify changes in risk, they can affect the direction of prices, and most important they can be exploited for profit. The methods for finding them are simple, and made more so by the use of a spreadsheet program or statistical software. These will be discussed in this chapter along with some practical applications.

THE SEASONAL PATTERN

Seasonal patterns are easier to find than the longer-term cycles or economic trends, because they must repeat each calendar year. Although any 12-month period can be used to find the seasonal pattern, academic studies in agriculture usually begin with the new crop year for grains, just following harvest, when prices tend to be lowest. This approach makes the carrying charges, which increase steadily throughout the new crop year, more apparent. For uniformity, the examples in this chapter always begin with a calendar year, which assumes no knowledge of where the season starts, and can be equally applied to stocks. Carrying charges are always reflected in the market price.

United States agricultural production is considered to be the standard for "seasonal," even though a wheat crop is harvested continuously throughout the year in different parts of the world. Prices are expected to be lower during the U.S. harvest and highest during the middle of the growing season. The influence of world stocks and anticipated harvest from other major producers in South America or Russia will cause an overall dampening or inflating of prices, rather than change the seasonal pattern. There is a constant flow of agricultural and industrial products throughout the world. The seller of a product will always choose the highest price as denominated in the local currency, so that a buyer with a weaker currency will appear to pay more than one with a stronger currency. But the seller gets the same price from everyone, net of shipping costs. The fact that any buyer can go to any seller creates the competition that produces a *world market price*. The interchangability of product is called *fungibility*.

Industrial commodities have seasonal price variation based on demand. Silver and gold, although increasingly used in electronics as high-end conductors, are still mostly consumed for jewelry but during unstable economic times, they serve as a hedge against inflation by the general public. Almost half of all copper is used in electrical and heat conductivity, with much of it in the form of an alloy with nickel and silver. Its seasonality is heavily related to the housing industry, where it is required for both electrical and water systems. New sources of ore are introduced infrequently, and the possibility of discovery or expansion is rarely seen in price movement as short-term anticipation. The primary supply problems in copper are related to labor as well as social and political changes in producing countries. More recently, growth in China and India has spurred a demand for copper. It is said that China has warehoused large quantities of copper in anticipation of future needs. This might distort the seasonal patterns, but they will return to normal as seasonal consumption becomes the driver.

There are many businesses with finished products that have seasonal demand, and their publicly traded stock prices will reflect that tendency. Because the shares in a company are far removed from buying and selling the raw materials that they use, changes in the price of raw materials may have a small effect on the bottom line, or the share price. Yet some industries, such as airlines, still show seasonal patterns, and the same procedures given here can be used to find them.

POPULAR METHODS FOR CALCULATING SEASONALITY

Seasonal patterns are most often calculated using monthly data, although some studies have tried to pinpoint their periodic turns to specific days. As with most other analysis, closer observation and shorter time periods also bring more noise and erratic results. With this in mind, the seasonal studies in this chapter will use monthly data and keep an eye toward the big picture.

An Important Note about the Data

There is one important caveat about the futures prices used in some of these examples: They were created using continuous, back-adjusted data. Continuous futures data are much easier to study than many individual delivery months and more accessible than cash data for most traders. Patterns in futures data will be more relevant for trading because it is the same data that is traded. Unfortunately, back-adjusted futures do not give the same results as using cash prices, and it is the cash price pattern that we normally see as the traditional seasonal pattern. This chapter will include comparisons of seasonal patterns using different types of data.

To create a continuous data series from individual contract delivery months, it is necessary to close the price gap between the old futures contract and the new one on the day of the roll and then adjust all the data backwards to reflect the value of that closed gap. The continuous data works well for trend analysis, but the oldest data may be much higher or lower than actual prices at that time due to the aggregate effect of adjusting dozens of gaps. In the case of wheat, the price in 1965 was $1.72/bushel using cash data but when back-adjusted it is $12.45. In other cases, such as interest rates, prices that have been adjusted back 25 years can actually become negative because the carry is always negative. A traditional seasonal analysis will use cash prices, rather than a constructed series, in order to have valid percentages. When using continuous futures data, results cannot be expressed as percentage changes. Cash prices for the underlying futures contracts are available on downloading services, such as Commodity Systems, Inc. (CSI), Boca Raton, Florida.

You should also remember that the price of a stock may have changed due to splits or reverse splits; therefore, the percent values we get for data in 1975 may not have been the same value at the time.

Seasonal Decomposition

Before showing the various ways of finding seasonal patterns manually, there are more automated ways of finding many of the same results. Statistical software will have options for *seasonal differences* and *seasonal decomposition*. For seasonal differences you are asked to enter the number of days or periods in the season. For monthly data that would be 12 and for daily data, probably 252 (you will want to divide the total number of days in the data by the number of years to get the exact number). The result will be a detrended price series, comparable to what will be done in the next section.

Seasonal decomposition is more extensive and essentially separates the main components of a time series, the trend, seasonal, cycle, and noise. It should provide the residuals; that is, the price fluctuations due to seasonality, after the other components have been extracted. In this chapter we will not be as rigorous about these other time series components, but try to be practical about extracting seasonal patterns that can be traded.

Detrending

It was already stated that the price trend, or the underlying change in currency, can make it more difficult to recognize the seasonal patterns. As an example, Figure 10.1 shows monthly average cash wheat prices from 1985 through 2010. There are both an upwards trend and an increase in volatility over time. By performing a simple regression analysis using Excel, then creating the regression line based on the slope and *y*-intercept, it is easy to see the long-term trend of prices, increasing about ½ cent per bushel each month from 1985. By subtracting the regression line values from the actual price we get the *residuals*, which are the *detrended wheat prices* also plotted in Figure 10.1. Volatility in 2007 is actually higher in the detrended prices than the cash prices.

It would be more accurate to remove the trend from all data before finding the seasonal patterns, but then the trend bias would need to be added back to find the net move for a specific month. Most traders will not want to do that, and in many cases, such as wheat, the underlying trend is a modest 0.42% per year and 51% over 25 years. Because this repeats equally for each month, it will not change the seasonal patterns.

Basic Calculation Components

Average Prices

Finding the seasonal pattern does not need to be complicated; however, some basic rules must be followed. For most analysts it is easiest to begin with a spreadsheet, where
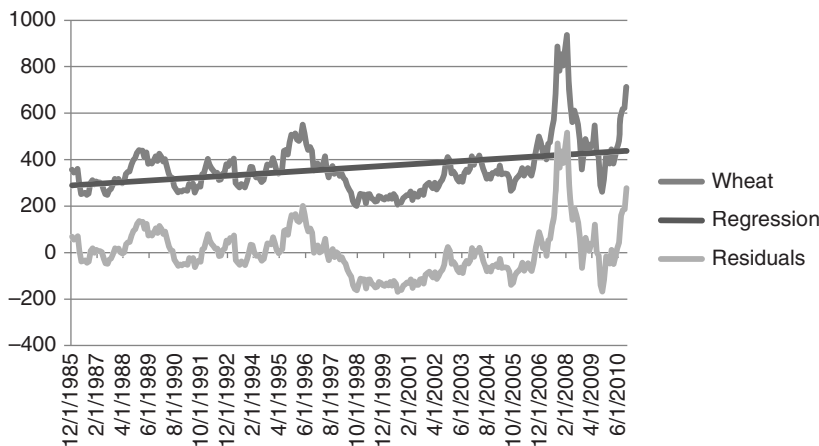
FIGURE 10.1 Detrending monthly cash wheat prices.

TABLE 10.1 Prices of Cash Wheat, 1966–2011, with Monthly Average and Median Values

(a) Cash Prices

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1966					1.72	1.86	1.89	1.88	1.74	1.72	1.79	1.76
1967	1.61	1.78	1.79	1.66	1.66	1.49	1.52	1.46	1.53	1.50	1.44	1.45
1968	1.49	1.48	1.49	1.36	1.37	1.26	1.26	1.19	1.19	1.30	1.32	1.39
1969	1.38	1.32	1.29	1.32	1.28	1.25	1.28	1.29	1.33	1.36	1.44	1.51
1970	1.48	1.57	1.51	1.58	1.42	1.43	1.46	1.57	1.70	1.77	1.74	1.75
...												
2007	4.33	4.50	4.01	4.68	4.75	5.34	5.71	6.84	8.87	7.83	8.54	8.05
2008	8.68	9.36	7.37	6.31	5.62	6.11	5.84	5.53	4.85	3.58	4.24	4.89
2009	4.60	4.33	4.66	4.63	5.47	4.31	4.15	2.89	2.62	3.14	4.13	4.10
2010	3.84	4.44	3.83	4.16	4.52	4.77	5.94	6.18	6.22	7.13	6.73	7.81
2011	8.37	8.03										
Average	3.553	3.526	3.324	3.242	3.137	3.117	3.176	3.219	3.315	3.394	3.548	3.574
Median	3.495	3.480	3.385	3.325	3.245	3.175	3.160	3.145	3.180	3.430	3.430	3.565
Adj Avg	0.062	0.054	−0.006	−0.030	−0.062	−0.068	−0.050	−0.037	−0.009	0.015	0.061	0.069
Adj Mdn	0.048	0.044	0.015	−0.003	−0.027	−0.048	−0.052	−0.057	−0.046	0.029	0.029	0.069

(b) Back-Adjusted Futures Prices

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1966					12.45	12.63	12.64	12.66	12.49	12.44	12.51	12.43
1967	12.32	12.42	12.42	12.29	12.32	12.18	12.16	12.06	12.11	12.07	11.99	11.97
1968	11.96	11.96	11.94	11.80	11.81	11.70	11.64	11.56	11.52	11.63	11.62	11.62
1969	11.62	11.58	11.51	11.55	11.54	11.51	11.46	11.50	11.51	11.52	11.57	11.60
1970	11.57	11.64	11.60	11.67	11.59	11.66	11.68	11.80	11.85	11.90	11.88	11.83
...												
2007	7.97	7.98	7.82	8.09	8.20	9.14	9.29	10.53	12.01	10.81	11.27	11.24
2008	11.64	13.68	13.00	11.18	10.54	12.04	10.82	10.80	9.37	8.09	8.12	8.51
2009	8.27	7.60	7.49	7.54	8.53	7.52	7.09	6.61	6.14	6.72	7.19	6.90
2010	6.33	6.35	6.03	6.13	5.87	5.58	7.28	7.74	7.58	7.92	7.24	8.18
2011	8.56	8.20										
Average	11.27	11.28	11.28	11.23	11.25	11.25	11.26	11.34	11.36	11.33	11.31	11.31
Median	11.62	11.63	11.55	11.46	11.43	11.51	11.30	11.50	11.51	11.40	11.57	11.60
Adj Avg	−0.0016	−0.0010	−0.0005	−0.0054	−0.0034	−0.0032	−0.0026	0.0043	0.0060	0.0037	0.0020	0.0016
Adj Mdn	0.0096	0.0105	0.0035	−0.0040	−0.0063	0.0007	−0.0182	−0.0002	0.0007	−0.0095	0.0053	0.0079

the months are recorded in each column and the rows represent years (see Table 10.1). The average monthly price, placed in each cell, can give a good indication of seasonal patterns by simply averaging each column and plotting the results as shown in the four summary lines at the bottom. The major criticism of this technique is that it ignores the changing price levels over time. For example, a 25-year study of soybeans will use prices that vary from $6 to $15 per bushel; price changes at the $15 level, which will be much more volatile, could overwhelm other years.

Getting the monthly averages can be time consuming, but with many charting and data services you can convert a daily chart to monthly (showing the last price of the month rather than the average) and then download those prices into an Excel spreadsheet. In the case of TradeStation you can write a program to print the prices shown on a monthly chart into a flat file.

Comparing Cash and Futures Seasonal Calculations

To show how important it is to use the correct data, the monthly last price of cash and back-adjusted futures for wheat was put into a spreadsheet (the entire spreadsheet is available on the Companion Website as *TSM Cash and Futures monthly tables wheat*) shown in Table 10.1*a* and *b*. In particular, the cash prices progress steadily from lower to higher, while the futures prices begin at higher prices and end at lower prices. There were periods of higher and lower volatility not shown in these tables.

At the bottom of each table are the monthly averages and monthly medians. Normally, the median is considered a better gauge of the typical price because it is not affected by one or two extremes, but shows the price that would normally be received. If we average all the average monthly prices, then divide each monthly average by that yearly average, we get the adjusted average that is shown two lines below. The lower two lines show the averages and medians divided by the yearly average (*Adj Avg* and *Adj Mdn*). Plotting the ratios for the averages and medians gives a picture of the seasonality in Figures 10.2a and b.

While the cash prices show a classic seasonal pattern, the futures do not. Winter wheat, the biggest U.S. crop, is harvested beginning in June and may continue through the summer. Figure 10.2a shows the lowest average prices in May and June, and the lowest median prices in July and August. Both are reasonable. The futures prices show an exaggerated picture, with the median prices spiking down in July but the average prices out-of-phase with what is believed to be the correct seasonal pattern. The conclusion is that futures prices may work for some markets, but cash prices are safer.

A Program to Create a Table of Monthly Price Changes

Creating a table of price changes in order to evaluate seasonality is a tedious task. The monthly changes, shown in various tables, were created using three TradeStation programs, *TSM Seasonal Average*, *TSM Seasonal Median*, and *TSM Seasonal Volatility*,
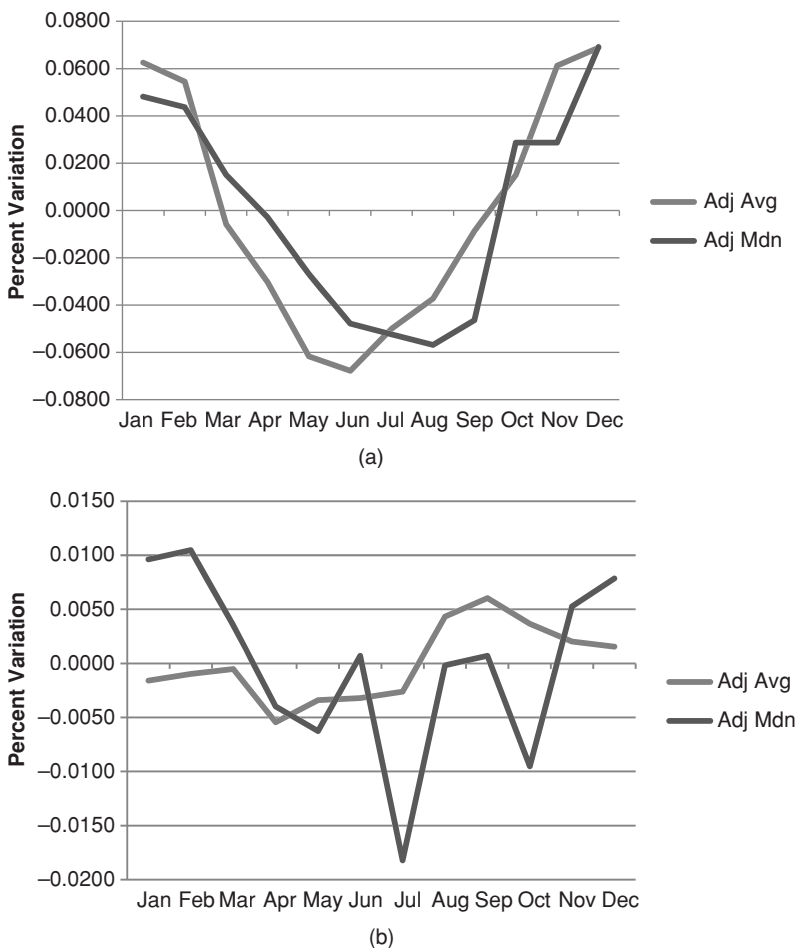
FIGURE 10.2   Seasonal patterns resulting from dividing the monthly averages by the average of all months. (a) The results using cash wheat prices. (b) The results using back-adjusted wheat futures.

available on the Companion Website. These programs write flat files in table form with the months along the top and the years along the left. The programs have no inputs. They will automatically run and produce the output table as soon as it is loaded into the chart workspace.

Indexing the Data

A simple way to adjust for price differences over time is by indexing data, where each new entry is based on a percentage change from the previous value. This method will work well for seasonal studies, but must use unadjusted cash data or stocks. Back-adjusted futures cannot be used for this purpose. Begin by giving the first monthly average price, for example, $25, the index value of 100. Each subsequent monthly average

price increases or decreases the index by the percentage change. If the second month had an average price of $27, then $100 \times 27/25 = 108$, an increase of 8%. If the third month showed an average of $26 then the index value becomes $108 \times 26/27 = 104$. This method can be found in the section "The Index" in Chapter 2. Once the data has been indexed, the average results are interpreted as percentage changes, which should be more useful as prices vary over time.

Consider two stocks, Southwest Airlines (LUV) and Amazon (AMZN). We would expect Southwest to have a seasonal pattern because most leisure travel occurs during the summer. Amazon is not as clear. First, the seasonal values are calculated from prices in the same manner as wheat, using data from August 1998 through February 2011. Results are shown in Table 10.2a with the adjusted average along the bottom. The prices are then indexed, starting at 100, and the months averaged in the same way. Figure 10.3 compares the seasonality in terms of price with those in terms of percent. Table 10.2a Southwest Airlines seasonality calculated in dollars ($) in part a and in percent (%) in part b.

In Figure 10.3 both seasonal patterns are the same, but the one expressed in dollars ($) appears to have very small fluctuations. Table 10.2a shows that the seasonal range is from about –$0.50 to +$0.50. The percentage chart is much clearer, showing fluctuations of –6% to +6%. Because Southwest remained at about the same price level for the entire period, the chart actually differs only in the scale. Most important, it shows that there is a seasonal peak from April through July (spring through early summer in the United States) and another in November for Thanksgiving holiday. August, September, and January are lows when families get back to work and school and after the major holidays.

The same seasonal calculations were done for Amazon, without any expectation of whether there was seasonality. Figure 10.4 shows the results; however, instead of showing the price and percentage fluctuations on the same scale, the percentage change is shown on the left and the price on the right. The charts are nearly identical, but the percentages show variations of –40% to +30% while the prices ranged from –$5.00 to
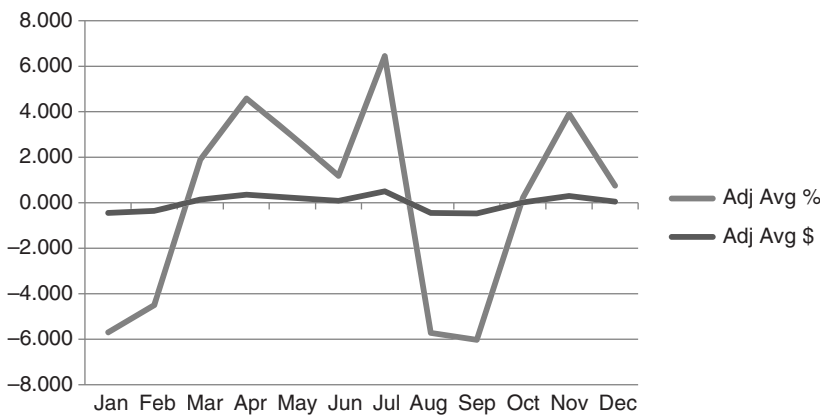
FIGURE 10.3 Seasonality of Southwest Airlines in dollars ($) and percent (%).

TABLE 10.2 Southwest Airlines Seasonality

(a) In Dollars

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1998								7.78	8.84	9.22	9.47	9.91
1999	11.77	13.16	13.22	14.23	14.01	13.60	12.13	10.94	9.96	11.02	10.70	10.58
2000	10.45	12.09	13.66	14.23	12.59	12.43	15.51	14.86	15.92	18.71	20.72	22.02
2001	20.57	18.33	17.49	17.94	19.71	18.22	19.72	17.64	14.63	15.67	18.48	18.22
2002	18.68	20.82	19.08	17.96	16.80	15.94	13.62	14.02	12.89	14.41	16.48	13.72
2003	12.88	11.92	14.18	15.76	15.87	16.99	16.21	16.89	17.49	19.17	17.77	15.95
2004	14.78	13.65	14.05	14.12	15.34	16.59	14.31	14.66	13.48	15.60	15.56	16.11
2005	14.33	13.71	14.10	14.73	14.41	13.80	14.05	13.20	14.71	15.86	16.35	16.28
2006	16.31	16.63	17.84	16.08	15.96	16.23	17.84	17.18	16.53	14.91	15.58	15.20
2007	14.98	15.01	14.59	14.24	14.21	14.81	15.55	15.01	14.70	14.12	14.06	12.12
2008	11.65	12.19	12.33	13.16	12.98	12.97	15.50	15.15	14.43	11.72	8.60	8.58
2009	7.00	5.87	6.31	6.95	6.71	6.71	7.83	8.15	9.58	8.38	9.18	11.41
2010	11.31	12.55	13.20	13.16	12.42	11.10	12.03	11.04	13.06	13.75	13.31	12.97
2011	11.84	11.83										
Average	13.58	13.67	14.17	14.38	14.25	14.12	14.53	13.58	13.56	14.04	14.33	14.08
Adj Avg $	−0.443	−0.350	0.147	0.356	0.227	0.092	0.501	−0.44 5	−0.46 8	0.018	0.304	0.059

(b) In Percent

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1998								100	113.625	118.509	121.722	127.378
1999	151.285	169.152	169.923	182.905	180.077	174.807	155.913	140.617	128.021	141.645	137.532	135.990
2010	145.373	161.311	169.666	169.152	159.640	142.674	154.627	141.902	167.866	176.735	171.080	166.710
2011	152.185	152.057										
Average	174.56	175.76	182.14	184.83	183.17	181.44	186.70	174.53	174.23	180.48	184.16	181.01
Adj Avg %	−5.691	−4.495	1.893	4.582	2.921	1.186	6.445	−5.721	−6.017	0.231	3.909	0.755
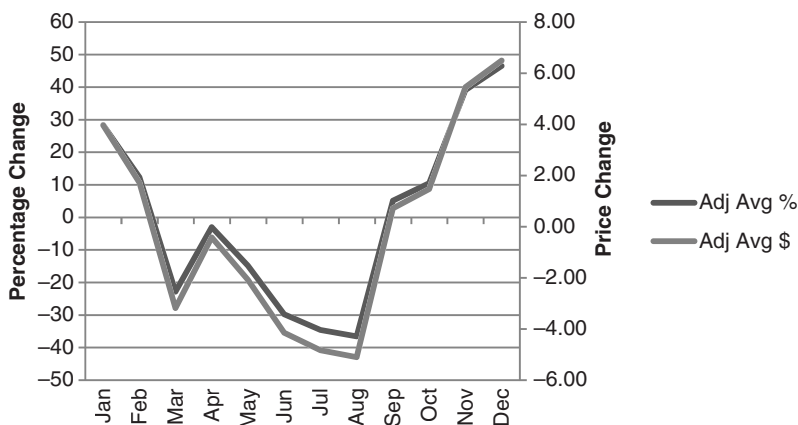
FIGURE 10.4 Amazon seasonality, using both price change and percentage change, August 1998 through February 2011.

+$6.50. During this period, the price of Amazon was as low as $6 and as high as $180. The average price fluctuation will not be a good measure of expectation when prices are at either extreme. Percentages will be better, but not perfect. Other techniques will be considered to see if the pattern changes, rather than the magnitude of the move.

The seasonal pattern for Amazon is remarkably clear, with business improving after the summer holidays and spiking ahead of the Christmas. In retrospect, retailers get the bulk of income at Christmas, so this result can be justified. Most important, this clear pattern could help trading decisions.

Detrending Amazon

Detrending, as seen in the previous section, is theoretically essential to finding a clear seasonal pattern. Although seasonality will always affect prices, a strong trend can overwhelm the seasonal movements and make the patterns invalid. Detrending can be accomplished by using linear regression, moving averages, yearly averages, and first differences, as well as the more complicated techniques of *link relatives* and X-11. As an example of the effect of detrending, consider Amazon, which has had a wide price range since 1998.

To detrend Amazon, the same steps were used as before:

1. Create a column with the simple numeric sequence, 1, 2, 3, . . . This is needed as the independent variable, x, in the regression because the date form does not work and the date skips weekends, which would cause incorrect results.

2. Find the slope and y-intercept, using the numeric sequence as x and the prices as y. Use the regression tool in Excel found in Data Analysis.

3. Create the linear regression line value for each corresponding price, x, using the formula $y = ax+b$, where a is the slope, b is the y-intercept.
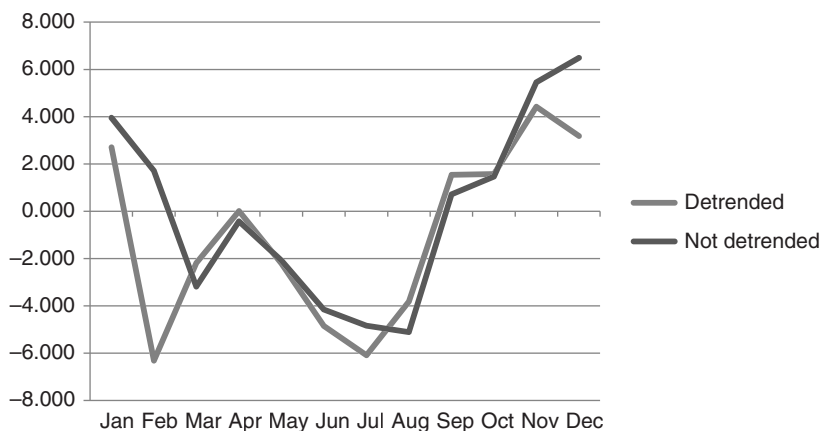
FIGURE 10.5 Comparison of seasonal patterns using original Amazon prices and detrended prices.

4. Calculate the residuals by subtracting each value on the regression line (3) from the corresponding original price.

5. Put the residuals into a table with the year on the left and the months across the top in order to create the seasonal chart.

When the detrended seasonal pattern is compared to the non-detrended pattern, as shown in Figure 10.5, the overall shape is the same. The major differences are that February is now much weaker and November is the high; November and December remain the two strongest months. Although this is only one case, we would conclude that, for finding the overall seasonal pattern for trading, it is not necessary to detrend prices. A trader selecting the strongest six months to bias trades on the long side would get the same results with original or detrended prices.

Removing the Trend with First Differences

Jake Bernstein, most well-known for his seasonal studies,[1] uses the method of first differences to remove the trend from prices before calculating the seasonal adjustment factor. He offers the following steps for determining the cash price seasonality:

1. Arrange the data used in a table with each row as one year. Columns can be daily, weekly, or monthly although most analyses will use monthly. Average prices are preferred for each period (see Table 10.2).

2. Compute a second table of month-to-month differences by subtracting month 1 from month 2, month 2 from month 3, and so on. This new table contains *detrended* values. The average of these differences should be near zero.

---

[1] Jacob Bernstein, *Seasonal Concepts in Futures Trading* (New York: John Wiley & Sons, 1986).

3. Calculate the sum of the price differences in each column (month) in the new table. Find the average for that column by dividing the number of years of data (columns may have different numbers of entries). This is the average price change for that month.

4. From the table, count the times during each month (column) that prices were up, down, or unchanged. This will give the frequency (expressed as a percent) of movement in each direction.

Bernstein adds the average monthly changes together, expresses the frequency of upwards price changes, and presents the results of corn in Figure 10.6.

Using Bernstein's method, the seasonal price tendency for corn is shown in Figure 10.6. While Figure 10.6 shows a clear pattern for corn, where monthly price changes fluctuate around the zero horizontal line, the same method would not work well for Amazon. Normally, the method of first differences will remove the trend; however, corn prices remained in a narrow price range for the 20 years while Amazon varied over a very wide range, starting low and ending high. The large changes in more recent years greatly distort the results.

	J-F	F-M	M-A	A-M	M-J	J-J	J-A	A-S	S-O	O-N	N-D	D-J
%	48	60	80	72	45	59	62	63	82	60	82	70
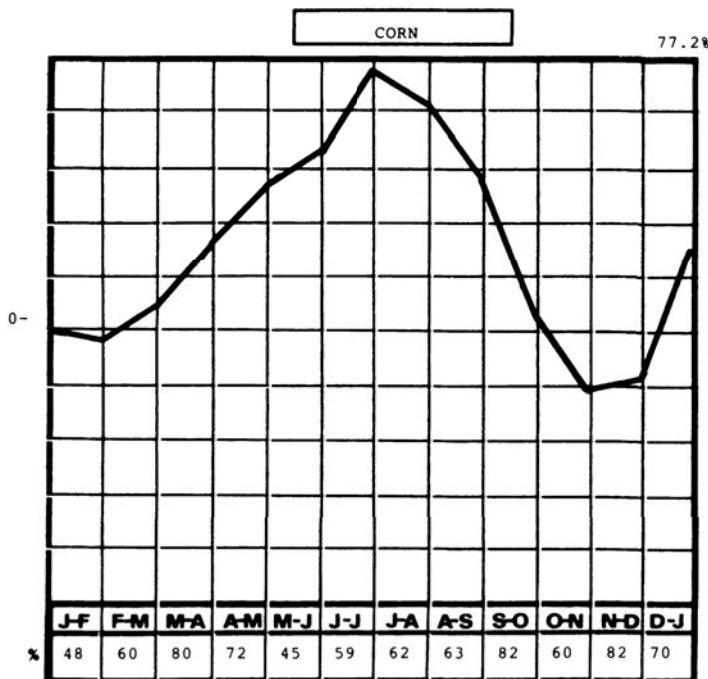
FIGURE 10.6 Seasonal price tendency in monthly cash average corn prices (1936 to 1983).

Effects of a Few Volatile Years

Seasonal calculations can be distorted by a few very volatile years that show large percentage changes, especially during a time when there is not normally a strong seasonal bias. For example, cotton has recently been the target of severe supply disruptions—flooding in Pakistan, geopolitical events in Egypt, and drought in the United States. The normal seasonal pattern based on the years 1973 through February 2011, shown in Figure 10.7, is similar to other northern hemisphere summer crops, with seasonal lows at harvest, or anticipation of harvest, July through September.

If only the past two years are averaged, the pattern is very different, with a price jump of 17% in October and rising prices during what is normally harvest season. This distortion points out that seasonal patterns are formed over many years, and a few nonseasonal events do not affect the long-term pattern—although they could result in large trading losses. It is interesting to note that these nonseasonal extremes are usually in the direction of higher prices.

Another way of seeing extremes is to plot the highest and lowest price changes that occur during any month. This also relates to volatility, which is discussed in a later section of this chapter. Although seasonal lows are in July, August, and September, the maximum monthly move was 45% in September, as seen in Figure 10.8, with the second largest move in August. The largest downside move of about 22% was in October. The upward surprises are much larger than the downward ones.

If the monthly percentage changes are each sorted highest to lowest over all years, the extent of the extremes can be seen in Figure 10.9. Although it is difficult to distinguish one month from another, the largest upside move was 45% in September, followed by 40% in August and nearly 30% in June. On the downside, July, September, and October all compete with each other at about –20%. Harvest seems to be the time for the largest moves, regardless of direction. Complete seasonal tables in prices and
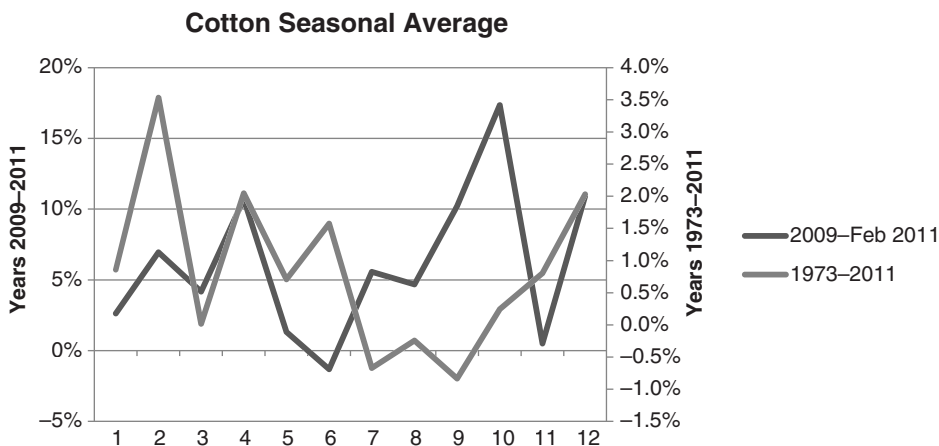
FIGURE 10.7   Cotton seasonal pattern from monthly averages, 1973 to 2011.
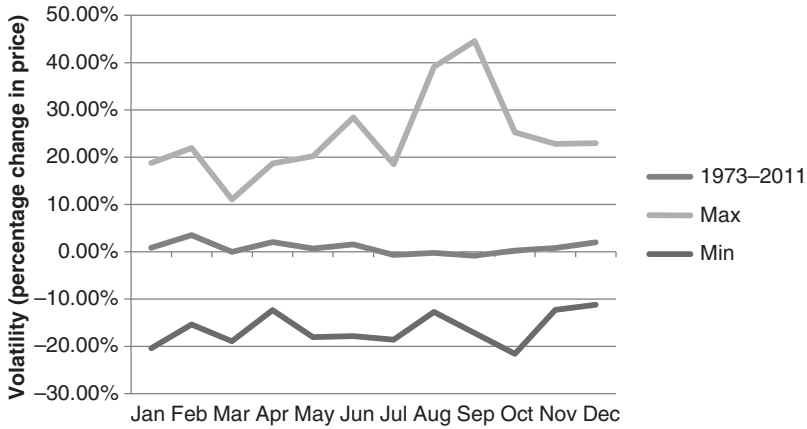
FIGURE 10.8 Monthly extreme moves in cotton, 1973 to February 2011.

percentages are available on the Companion Website as *Cotton Extreme Years* and *Heating Oil Extreme Years*.

Method of Yearly Averages

Although a number of different methods have been used to show the seasonal patterns, all of them have problems when prices have wide ranges or have been inflated. Expressing prices as percentages tends to reduce some of the issues, but changes at higher price levels still seem to overwhelm those at lower levels. Analysts have been trying to correct this problem for years.

An alternative way of expressing percentages, and one that could be used with back-adjusted futures, is to express each year's variation as it relates to the annual average.
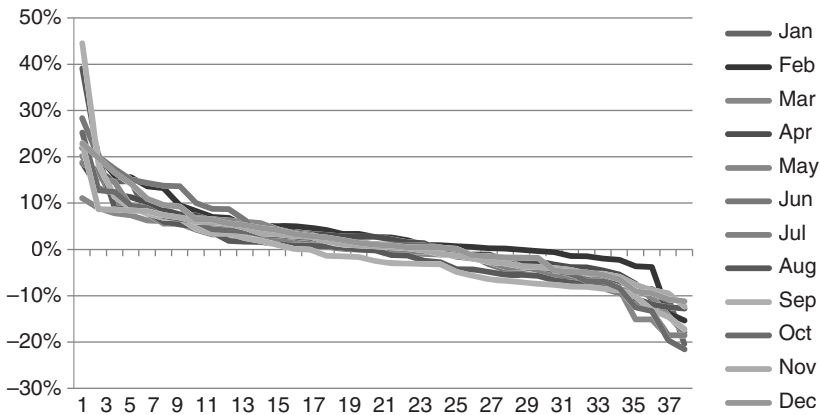
FIGURE 10.9 Cotton percentage monthly changes, sort highest to lowest.

TABLE 10.3a Original Cash Wheat Data with Yearly Averages

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Average
1966					172.00	185.50	189.25	188.00	173.75	172.00	179.00	176.00	179.44
1967	161.00	178.25	179.25	166.00	166.25	149.25	151.75	146.25	153.00	149.50	144.25	145.00	157.48
1968	149.25	147.50	148.50	135.50	136.75	125.50	125.50	118.50	119.25	129.75	131.50	138.50	133.83
1969	137.50	132.00	128.75	132.25	128.25	124.75	127.75	129.25	132.75	136.25	144.00	150.50	133.67
1970	148.00	156.50	150.50	158.00	141.75	143.00	146.25	156.50	169.75	176.50	174.25	174.50	157.96

TABLE 10.3b Monthly Percentage Changes Based on the Yearly Averages of the Same Year

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1966					-0.041	0.034	0.055	0.048	-0.032	-0.041	-0.002	-0.019
1967	0.022	0.132	0.138	0.054	0.056	-0.052	-0.036	-0.071	-0.028	-0.051	-0.084	-0.079
1968	0.115	0.102	0.110	0.012	0.022	-0.062	-0.062	-0.115	-0.109	-0.031	-0.017	0.035
1969	0.029	-0.012	-0.037	-0.011	-0.041	-0.067	-0.044	-0.033	-0.007	0.019	0.077	0.126
1970	-0.063	-0.009	-0.047	0.000	-0.103	-0.095	-0.074	-0.009	0.075	0.117	0.103	0.105

For agricultural products, using the crop year is better than the calendar year, but in the end the results should be the same.

In order to compare methods, the same wheat data was used as in Table 10.1. The first few years of the original data and the yearly averages are shown in Table 10.3a. The percentage changes for each month are calculated using the average of the same year, *change = month/average* – 1, and shown in Table 10.3b. The same process as before is followed to create the seasonal patterns. Each month is averaged down, then divided by the average of all months in order to see the results as variation around zero. Figure 10.10 shows that the pattern based on monthly averages is nearly identical to the first method of finding seasonality, but the median shows lows in May and June rather than August and September, a somewhat more consistent result.

A general formula for the monthly average is:[2]

$$APP_i = 100\left\{\frac{1}{N}\left[12\sum_{n=1}^{N}\left(P_{in} / \sum_{j=1}^{12}P_{jn}\right)\right]\right\}, i = 1,12$$

Where $APPi$ = the Average Percentage Price in month i
  i = the calendar month from 1 to 12
  N = the number of years in the analysis
  P_{jn} = the average price for month j of year n

This formula may be applied to weekly or quarterly average prices by changing the 12 to 52 or 4, respectively.
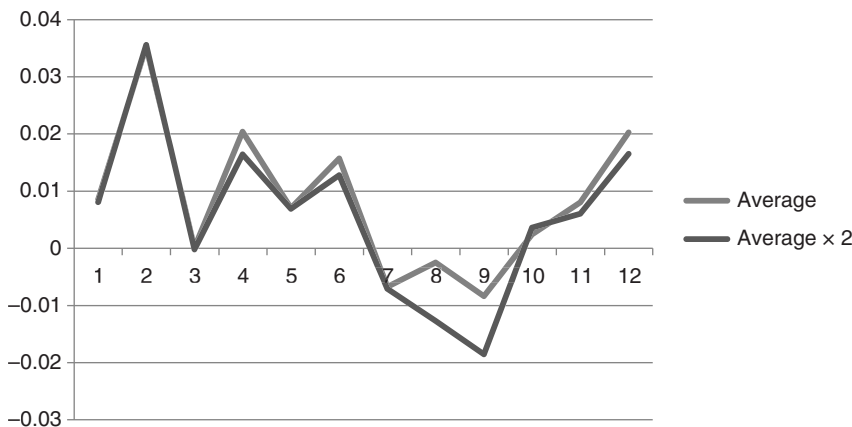
FIGURE 10.10 Season pattern of wheat using the method of yearly averages.

[2]David Handmaker, "Low-Frequency Filters for Seasonal Analysis," in Perry J. Kaufman, *Handbook of Futures Markets* (New York: John Wiley & Sons, 1984).

The use of an annual average price still does not account for a long-term trend in the price of the stock or futures market. If the rate of inflation in the United States is 6%, there will be a tendency for a commodity price, such as gold, to be 0.5% higher each month, resulting in a trend toward higher prices at the end of the year. Longer trends, such as the steady rise in grain prices from 1972 to 1975 and from 2008 to 2011, alternating with even longer declines, can obscure or even distort the seasonality unless the trend is removed. Seasonality is still a major factor influencing price variation, even when there is a dominant trend.

The Method of Link Relatives

Another technique for identifying seasonal price variations and separating them from other price components involves the use of *link relatives*. Table 10.4 shows the following steps for the first few years of wheat cash prices, 1966–1968, although the average and other calculations include the years 1966–2010. The full spreadsheet, along with all calculations, is available on the Companion Website as *TSM Wheat method of yearly averages and link relatives*.

1. Each month is expressed as a ratio by dividing each monthly price by the preceding one. December 1966 is used to find the ratio for January 1967, but then we ignore the 1966 data and begin in 1967 in order to have only those years with data in every month.

2. Find the average of the monthly ratios for all years. Put that value in the right column.

3. Find the average (or the median, which is preferred if an adequate sample is used) of the monthly ratios expressed in rows 1, 2, and 3. The average in row 4 represents monthly variation as a percentage of change. Thus far, this is the same as Bernstein's average monthly price changes, expressed as a percent of the prior price.

4. In order to establish a fixed base in the manner of an index, *chain relatives* (see Table 10.4b) are constructed always using January as 1.00; each monthly chain relative is calculated by multiplying its average link relative by the average link relative of the preceding month. In row 2 of part b, the March 1967 chain relative is then 1.107 (the February 1967 chain relative) × 1.006 (the March 1967 link relative) = 1.113 (the March 1967 chain relative).

5. A constant trend throughout the test period can be found by multiplying the December 1968 chain relative (part c) by the January 1967 link relative (part b) giving 0.928 × 0.915 = 0.849. If prices show no tendency for either upward or downward movement, the result would be 1.00 (the right column); however, long-term inflation should cause an upward bias and therefore the results are expected to be higher. This leaves a negative factor of 15.1% unaccounted for, indicating that the 1967–1968 years showed a strong downward bias; therefore, the expected rate of inflation was offset by some other economic factor, such as the accumulation of grain stocks by the U.S. government or a much better than expected crop.

TABLE 10.4 Wheat Prices Expressed as Link Relatives

(a) Original Price Data

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1966					172.00	185.50	189.25	188.00	173.75	172.00	179.00	176.00
1967	161.00	178.25	179.25	166.00	166.25	149.25	151.75	146.25	153.00	149.50	144.25	145.00
1968	149.25	147.50	148.50	135.50	136.75	125.50	125.50	118.50	119.25	129.75	131.50	138.50

(b) Link Relatives

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Average
1967	0.915	1.107	1.006	0.926	1.002	0.898	1.017	0.964	1.046	0.977	0.965	1.005	0.986
1968	1.029	0.988	1.007	0.912	1.009	0.918	1.000	0.944	1.006	1.088	1.013	1.053	0.997
Average	0.990	0.993	0.990	0.980	0.980	0.989	1.018	1.013	1.027	1.040	1.047	1.010	
Median	0.993	0.987	0.981	0.987	0.968	0.983	1.008	1.013	1.024	1.048	1.036	1.013	

(c) Chain Relatives (only years with all data)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Growth
1967	100.000	110.714	111.335	103.106	103.261	92.702	94.255	90.839	95.031	92.857	89.596	90.062	−0.009
1968	100.000	98.827	99.497	90.787	91.625	84.087	84.087	79.397	79.899	86.935	88.107	92.797	−0.006

(d) Corrected Chain Relatives

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Average
1967	1.000	1.098	1.104	1.022	1.023	0.918	0.933	0.899	0.941	0.919	0.886	0.891	0.969
1968	1.000	0.981	0.988	0.901	0.909	0.834	0.834	0.787	0.792	0.862	0.874	0.921	0.890

(e) Index of Seasonal Variation

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1.004	1.000	0.986	0.966	0.949	0.935	0.953	0.975	1.007	1.041	1.086	1.098

6. The chain relatives must be corrected (part d) by adding the negative bias back into the values, using the same technique as in computing compound interest. For example, from 1967 to 1977, the Consumer Price Index increased from 100 to 175, a total of 75% in 10 years. To calculate the annual compounded growth rate for that period, apply the formula:

$$Compound\ rate\ of\ growth = \sqrt[N]{\frac{Ending\ value}{Starting\ value}} - 1$$

where N = the number of years or the number of periods over which the growth is compounded

$$R = \sqrt[10]{\frac{175}{100}} - 1$$
$$= 1.05755$$
$$= 0.05755$$

This indicates a compounded rate of inflation equal to 5.75% per year. In the case of wheat, if the trend had been positive, that is, greater than 1.00 instead of 0.990, the growth rate would be subtracted from each month to offset the upward bias. In this case, the results are added back into the chain relative to compensate for the negative influence. A 0.1% decline, compounded over 12 consecutive entries, gives:

$$R = \sqrt[12]{\frac{99}{1.00}} - 1$$
$$= 0.99916 - 1$$
$$= -0.00084$$

7. This is a compounded deflation of about 8/100 of 1%. To apply this to wheat, calculate the compounded rate of change for the net trend change each year. In 1968, the trend bias of 0.919 becomes a compounded growth rate of –0.0070. The growth rate for each year is added to each monthly chain relative of that year to get the corrected chain relative. For example, growth rate of –0.007 for 1968 is added to the February chain relative, 0.988, to give 0.981.

8. The chain relatives have been calculated on a base of January, which was important in order to correct the compounded bias throughout the test period. The final step is to switch the corrected chain relatives to a base of the average value. Using the corrected chain relatives, find the average of each month and then calculate the average of all the months. Divide the monthly averages by the yearly average to normalize the results. The final result is the *Index of Seasonal Variation*, shown in part *e*.

Figure 10.11 shows the monthly averages during the four main steps needed to create the index. The final index is very smooth and has its seasonal lows at harvest. A complete study of seasonality using this method can be found in Courtney Smith, *Seasonal Charts for Futures Traders*.
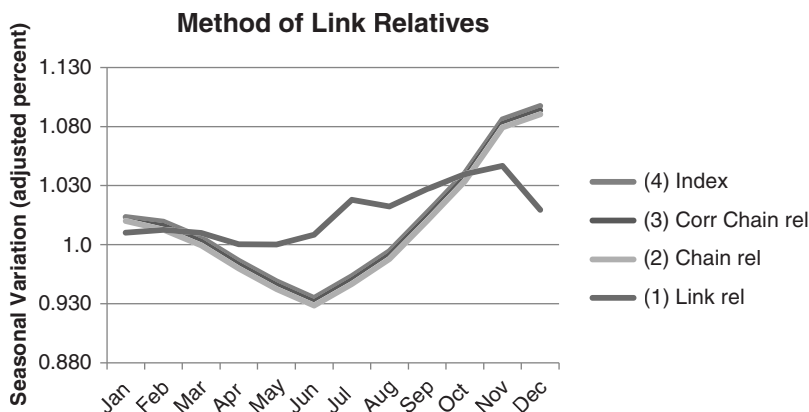
FIGURE 10.11 The four steps in creating link relatives.

The Moving Average Method

The moving average is a much simpler yet very good technique for determining seasonal patterns and removing the trend. Using monthly cash wheat prices, calculate as follows:

1. Take the average of the first 12 monthly prices, May 1966 through April 1967. Quarterly prices can also be used but not as desirable.
2. Place the average (1) in the sixth row of column 3, which is the same as lagging the average by 6 months.
3. In column 4 take the average of the sixth and seventh rows because a calendar year has an even number of months; therefore, the midpoint will be the average of month 6 and month 7.
4. Get the *seasonal adjustment factor* by subtracting the midpoint values (3) from the corresponding average price.
5. The *seasonal index* = 1 + seasonal adjustment factor/100.

These calculations can be seen in Table 10.5 and the midpoint plotted along with wheat prices from 2005 through 2010 in Figure 10.12. The midpoint shows the expected wheat price given seasonality, so that the difference between the actual wheat price and the midpoint is abnormal price movement.

The seasonal index is used to construct the familiar seasonal chart. By averaging all of the January seasonal index values, then all of the February, and so on, we can plot the seasonal pattern in Figure 10.13. The results are very smooth, easy to calculate, and are equally as good as the best results from any of the other methods.

X-11 and X-12 ARIMA Methods

The seasonal adjustment method *X*-11 (Census Method II-X-11) has been most widely used for creating a seasonally adjusted series of such information as car and housing

TABLE 10.5 Calculations for the Moving Average Method.

Date	Price	12-Month Moving Average, Lag 6	Midpoint	Seasonal Adjustment Factor	Seasonal Index
5/31/1966	172.0				
6/30/1966	185.5				
7/29/1966	189.3				
8/31/1966	188.0				
9/30/1966	173.8				
10/31/1966	172.0	176.67			
11/30/1966	179.0	176.19	176.43	−0.93	0.991
12/30/1966	176.0	173.17	174.68	2.82	1.028
1/31/1967	161.0	170.04	171.60	−3.10	0.969
2/28/1967	178.3	166.56	168.30	1.32	1.013
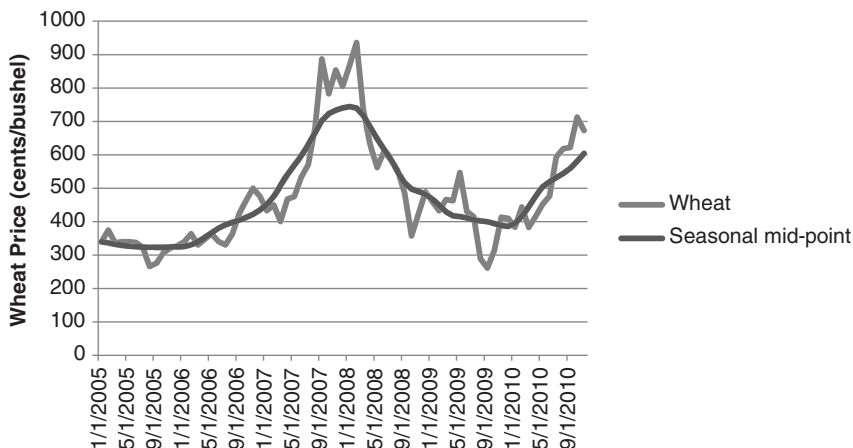
FIGURE 10.12 Wheat seasonal midpoint using the moving average method, 2005–2010.
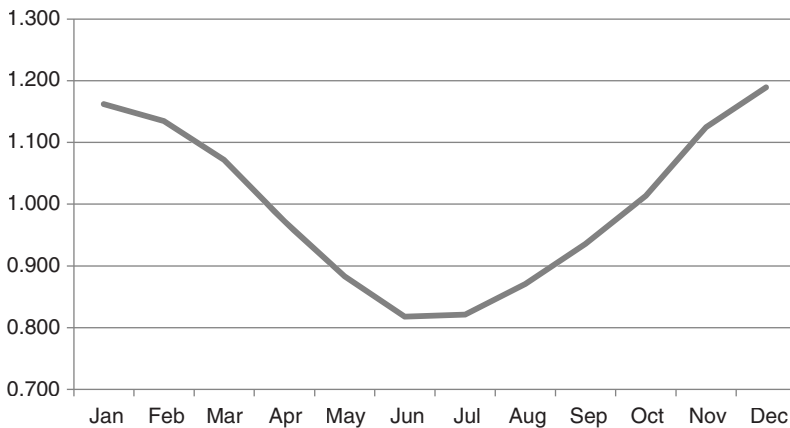
FIGURE 10.13 Averaging all of the monthly seasonal index values gives a standard seasonal pattern for wheat.

sales, as well as other consumer products. It is very thorough, using moving averages to detrend the data, and includes both an initial estimation and reestimation. *X*-12-*ARIMA*[3] is used to adjust the CPI and includes both X-11 estimations and a regression ARIMA calculation (see Chapter 6) for determining intervention and data extension. Because it is widely used by economists, an outline of the X-11 calculations follow.[4]

1. Calculate a centered 12-month moving average (MA). Subtract this MA from the original series to get an initial detrended seasonal series in the same manner as the moving average method.

2. Apply a weighted 5-period MA to *each month separately* to get an estimate of the seasonal factors.

3. Compute a centered 12-month MA of the seasonal factors (Step 2) for the entire series. Fill in the six missing values at either end by repeating the first and last available MA values. Adjust the seasonal factors from Step 2 by subtracting the centered 12-term MA. The *adjusted seasonal factors* will total approximately zero over any 12-month period.

4. Subtract the seasonal factor estimates (Step 3) from the initial detrended seasonal series (Step 1). This is the *irregular component series* used for outlier adjustment.

5. Adjust the outliers in Step 4 by the following procedure:

   a. Compute a 5-year moving standard deviation s of the irregular component series (Step 4).

   b. Assign weights to the series components c_i as follows:

   0 if $c_i > 2.5s$
   linearly scaled from 0 to 1 if $2.5s \geq c_i \geq 1.5s$
   1 if $c < 1.5s$

   Use this weighting function to adjust the detrended series in Step 1.

6. Apply a weighted 7-period MA to the adjusted series (Step 5) to get the *preliminary seasonal factor.*

7. Repeat Step 3 to standardize the seasonal factors.

8. Subtract the series resulting in Step 7 from the original series to find the *preliminary seasonally adjusted series.*

9. To get the trend estimate, apply a 9-, 13-, or 23-period *Henderson's weighted moving average*[5] to the seasonally adjusted series (Step 8). Subtract this series from the original data to find a second estimate of the detrended series.

---

[3]The U.S. Government provides details of its methods on the Bureau of Labor Statistics website. Scan the Internet for "X-11" or "X-12."

[4]A more detailed account of X-11 and Henderson's weighted moving average (Step 9) can be found in Abraham, Bovas, and Johannes Ledolter, *Statistical Methods for Forecasting* (New York: John Wiley & Sons, 1983),178–191. Their book also includes a computer program for "seasonal exponential smoothing."

[5]A specialized symmetric assignment of weighting values. A specific example can be found in Abraham, Bovas, and Johannes Ledolter, *Statistical Methods for Forecasting* (New York: John Wiley & Sons, 1983), 178.

10. Apply a weighted 7-period MA to each month separately to get a second estimate of the seasonal component.

11. Repeat Step 3 to standardize the *seasonal factors*.

12. Subtract the final seasonal factors from the original series to get the final *seasonally adjusted series*.

X-12-ARIMA

The Bureau of Labor Statistics uses intervention analysis in the seasonal adjustment of consumer price indexes to *provide more accurate CPI data*. This process offsets the effects that extreme price volatility would otherwise have on the estimates of seasonally adjusted data. It states,[6]

> *Intervention analysis is the prior adjustment of an index series before the calculation of the seasonal factors. Prior adjustment may be called for if a "ramp" occurs. (A "ramp" occurs where a good or service undergoes a unique, large, and rapid change in price level.) An example would be a large decrease in the price of gasoline due to the breakdown of an oil cartel. Removal of the ramp gives a clearer seasonal pattern and lessens the irregular component.*

It must be said that the BLS statement references a large drop in oil prices; however, it is of greater interest to them to reduce the impact of a large increase. Although they argue that oil prices are embedded in all other consumer prices, therefore, using them explicitly in an inflation calculation (such as the CPI) would be counting them twice, it is in the interest of the U.S. Government to show only small increases in the CPI. By law, the government must give cost-of-living increases to all Social Security beneficiaries based on the annual change in the CPI. A large upward revision would be very costly. Intervention analysis, although automated, provides a way to delay the effects of large shifts in prices, whether event-driven or a structural change in the seasonality. Because of this inherent conflict of interest, it becomes questionable whether the CPI is a good choice for deflating commodity prices.

In keeping with the ARIMA concept, each year the seasonal adjustment factors are recalculated using the most recent five to eight calendar years. These values are then used for the next year's seasonal adjustment indexes and the most recent five years of adjustment factors are published. As part of the process, it states,

> *Most higher level index series are adjusted by the indirect, or aggregative, method, which is more appropriate for broad categories whose component indexes show strongly different seasonal patterns. Under the aggregative method, direct*

---

[6]*BLS Handbook of Methods Bulletin 2490* (April 1997), Chapter 17, "The Consumer Price Index: Estimation of Price Change," 192.

adjustment is first applied to indexes at lower levels of detail, and thereafter the adjusted detail is aggregated up to yield the higher level seasonally adjusted indexes. If intervention analysis is indicated, it will be used in adjusting selected lower level indexes prior to aggregation. For those series which have not been selected for seasonal adjustment, the original, unadjusted data are used in the aggregation process.

Winter's Method

Another technique for forecasting prices with a seasonal component is *Winter's method*,[7] a self-generating, heuristic approach. It assumes that the only relevant characteristics of price movement are the trend and seasonal components, which are represented by the formula

$$X_t = (a + b_t) \times S_t + e_t$$

where

X_t = the estimated value at time t
$(a + bt)$ = a straight line that represents the trend
S_t = the seasonal weighting factor between 0 and 1
e_t = the error in the estimate at each point

If each season is represented by n data points, S_t repeats every n entries, and

$$\sum_{t=1}^{n} S_t = 1$$

The unique feature of Winter's model is that each new observation is used to correct the previous components a, b, and S_t. Without that feature, it would have no applicability to price forecasting of housing starts, employment, or other seasonal data. Starting with two or three years of price data, the yearly (seasonal) price average (for example, a 12-point average for monthly data) can be used to calculate both values a and b of the linear trend. Each subsequent year can be used to correct the equation $a + b_t$ using any regression analysis. Winter's method actually uses a technique similar to exponential smoothing to estimate the next a and b components individually. The seasonal adjustment factors are assigned by calculating the average variance from the linear component, expressed as a ratio, at each data point. As more observations are made, each component can be refined, and it will take on the form of the general long-term seasonal pattern.

---

[7] Winter's method, as well as other advanced models, can be found in Douglas C. Montgomery and Lynwood A. Johnson, *Forecasting and Time Series* (New York: McGraw-Hill, 1976), and Abraham and Ledolter, *Statistical Methods for Forecasting* (New York: John Wiley & Sons, 1983).

Changes in Seasonal Patterns

Very little remains the same over time. While the United States has improved its production methods and has always been a major exporter of wheat and other grains, the rest of the world improves as well. Wheat is grown in many countries and, along with soybeans and rice, is one of the important food staples. And while Russia has always been a big producer of wheat, South America has made the biggest steps forward.

In 1966 the world wheat production was about 300 million metric tons, of which the United States produced 35 million, China 25 million, and Argentina 6 million. In 2003 the world production was 552 million (up 84%), with the United States at 63 million (up 80%), China 86 million (up 244%), and Argentina at 13.5 million (up 125%). The estimated world production for the 2010–2011 crop year is only 351 million MT due to severe weather problems in many countries.

China is in the northern hemisphere; therefore, it has essentially the same crop-year as the United States and Russia. However, increased production in Argentina, Brazil, and Chile would affect our seasonal patterns. Countries needing wheat will purchase from any exporting country at the best available price. That makes wheat a fungible commodity. If the United States is in short supply of wheat, or asks too much because of inflation, then buyers go elsewhere, driving down the price of United States wheat until it reaches the world market value.

It may be too difficult to forecast how the seasonal pattern for wheat would change based on the production of other countries and the change in value of the U.S. dollar, but we can show how patterns have changed using historic prices. Figure 10.14 shows the seasonal pattern using monthly average prices (the first method discussed) for all years 1966–2010. The low price is clearly in June, with May the second-lowest price. If we look only at 1966, the low was in May; otherwise, the pattern is similar. More recently, 2001–2010, there is a shift backwards, showing the low in April with the second-lowest
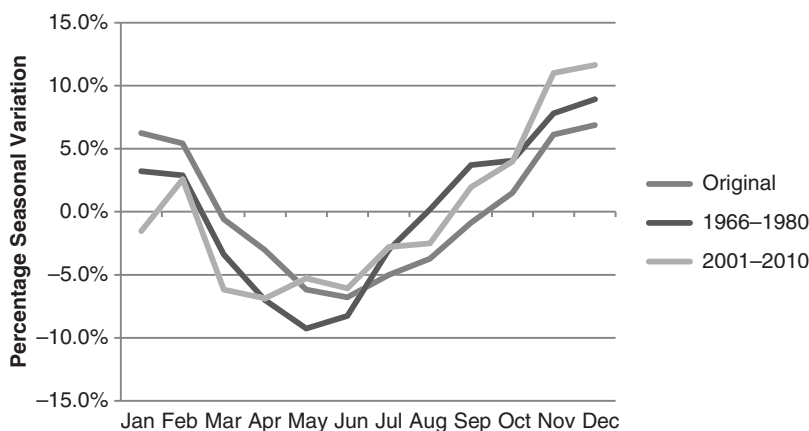
FIGURE 10.14 Changes in seasonal patterns for wheat.

month in March, and a somewhat more irregular pattern. These changes could be structural, based on changing supply and demand demographics, or they might be temporary due to the recent sharp drop in world production. Because this book is not intended to forecast fundamentals, it is only important that the seasonal pattern can change, and that trading systems that rely on seasonality should have a broader window when it comes to anticipating lower seasonal prices.

Seasonal Volatility

Consistent seasonal patterns can be confirmed by a corresponding increase in volatility as the season reaches its peak as seen in Figure 10.15, which shows heating oil from 1984 to 2010. Volatility was calculated first by finding the monthly percentage changes in price, then by averaging each month and finding the maximum and minimum volatility in that month. Based only on volatility, April, May, and June are the seasonal lows. This corresponds to the end of the winter heating season and the refining shift from heating oil to gasoline. During the summer months, there is a slow increase in inventories, but winter is the season of both high demand and high uncertainty. Unexpected cold spells, without considering geopolitical events, will draw down inventories and make heating oil prices more volatile.

Volatility tells us that December and January are the key seasonal months for heating oil. This can be important information for trading and may result in

1. Taking only long trades during the winter months
2. Reducing position size during the winter to offset the increase in risk
3. Reducing position size in May and June due to lower activity
4. Buying volatility using options.
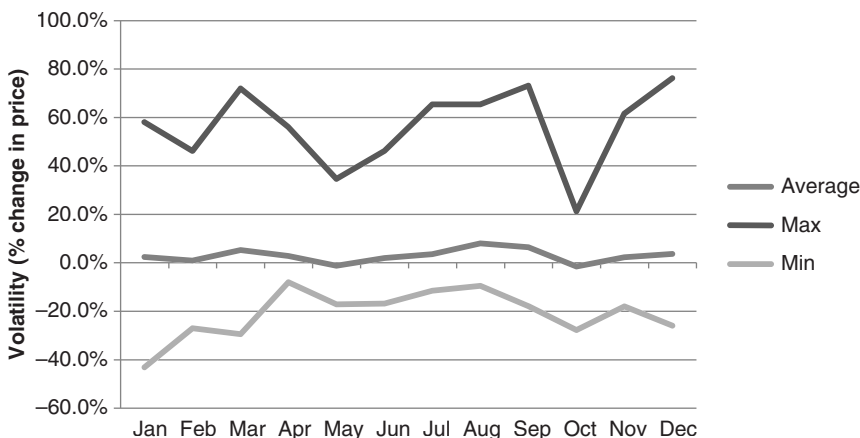
FIGURE 10.15 Heating oil monthly average volatility, 1984 to 2010, as a percentage of price.

For the purpose of confirming the validity of the seasonal results, the steady increase and decrease in volatility surrounding the peak month of June indicates a building of the seasonal concerns that cause wide fluctuations. If a single price shock had occurred unrelated to seasonality, there would be a sharp increase in volatility during one month, perhaps declining afterwards, but with no steady growth of volatility leading up to this event.

Weather Sensitivity

The effects of changes in weather, especially extremes in weather, are an inseparable part of the agricultural market's seasonal effects. Without weather, the price of an agricultural commodity would lack the surprises that cause them to jump around during the growing period. Once planted, you would have a very good idea of the expected supply. For insurance companies, hurricane season (from mid-summer through October) contains the same uncertainty.

Each agricultural product, and those firms impacted by changing agricultural prices, has its own particular sensitivity to weather. Grain planting can be delayed due to rain, causing some farmers to switch from corn to soybeans; hot weather during pollination will significantly reduce yields; a freeze in September can stop the ripening process and damage production for orange juice. Freezes are of greater concern than droughts and affect more products. Patterns of crop sensitivity to weather depend upon their location in the northern or southern hemisphere; however, active trade between world markets show that crops grown primarily in the northern hemisphere are affected by weather developments in the southern hemisphere. In Table 10.6 major weather events are

TABLE 10.6 Weather-Related Events in Southern and Northern Hemispheres

	Southern Hemisphere	Northern Hemisphere
January		OJ peak for Florida freeze, heating oil (cold or hot)
February	Corn pollination in S. Africa and Argentina, soybean pod development	Heating oil (cold or hot)
March		
April		Corn, cotton planting
May		Soybean planting
June		Cotton boll development
July	Coffee freeze Brazil, OJ freeze Brazil, Cocoa pod development in W. Africa, and pod rot in Brazil	Sugar: heat in Russia, corn pollination
August		Soybean pod development
September		Atlantic hurricanes affect sugar, orange juice, heating oil; corn freeze
October		Soybean freeze, cotton harvest
November	Brazil coffee rainy season	Soybean harvest corn harvest
December		Heating oil (cold or hot)
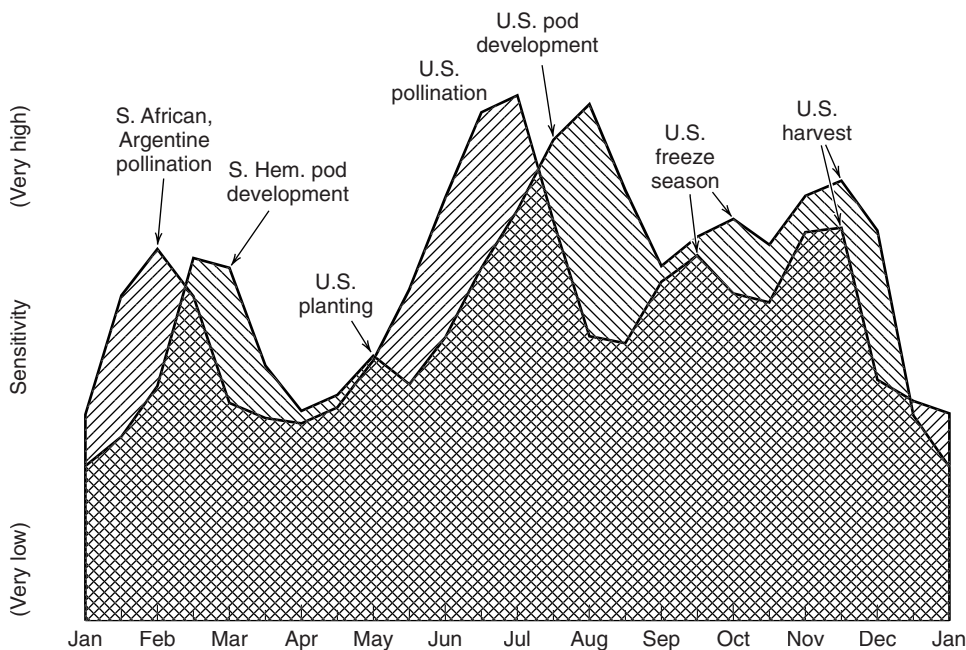
FIGURE 10.16   Weather sensitivity for soybeans and corn.
Source: Smith Barney.

separated into southern and northern hemisphere. Figure 10.16 combines the soybean and corn weather sensitivity charts[8] to show the effects of both southern and northern hemisphere weather.

Measuring Weather Sensitivity

While a weather sensitivity chart, such as the one shown in Figure 10.16, may appear to have a strong similarity to a standard seasonal chart, they are actually very different. Weather sensitivity is a measurement of potential price volatility. It could simply be the highest price that the market reached during a period in which the weather event occurred. Those more adept at statistics would want to record the increase in price relative to the average during those months in which weather was a factor and then show the price representing the 95% confidence level, about two standard deviations.

A thorough approach to measuring weather sensitivity is to record the temperature each day and find out whether it is unusually far above or below the average. Information available from the NOAA (National Oceanic and Atmospheric Administration) will give you both regional weather and measurements of effective heat, in thermals. It is the cumulative effect of heat that is damaging to a crop, rather than a high temperature for one

---

[8]Jon Davis, "Weather Sensitivity in the Markets" (Smith Barney, October 1994).

day. In addition, the amount of crop damage is not simply a linear relationship with rising temperatures but is more likely to start slowly and increase quickly once critical levels of time and heat are reached. The same is true of a drought. Advancements in hybrid seed production have made corn and other grains drought resistant. They can withstand long periods of dry weather. Many analysts have been unpleasantly surprised by recommending a long position in corn only to find that, even with low levels of rain, the harvest was a record crop.

Measuring the variations in temperature and its effect on crop output and price is done using *Heating Degree Days* (HDD) and *Cooling Degree Days* (CDD), where

$$HDD_t = \max(0, 65°F - \text{Average temperature}_t)$$
$$CDD_t = \max(0, \text{Average temperature}_t - 65°F)$$

The daily values are then accumulated during specific critical months. For example, heating oil would be December, January, and February, a total of 90 calendar days, where t is the number of days from December 1 but no greater than 90,

$$Cum\ HDD_t = \sum_{i=1}^{t} HDD_i$$

The demand for heating oil can be closely estimated by recording the temperature relative to population, so that a widespread hailstorm in August does not have the same impact in Montana as it does in Ohio. This has an immediate effect on futures prices, and a ripple effect on related businesses. During sustained cold periods, increased production of heating oil will mean decreased production of gasoline, driving those prices higher.

The Chicago Mercantile Exchange (CME) has made using heating degree days and cooling degree days easier by initiating futures contracts called *U.S. Monthly Weather Cooling Degree Day futures*, and the equivalent for heating degree days. In addition, there is a *Seasonal Strip Cooling Degree Day futures* contract (and heating degree days), which accumulates the next n delivery months of individual cooling degree day futures.

SEASONAL FILTERS

Seasonal filters identify periods during the year in which a long or short position would be favored based upon a clear seasonal pattern. These periods should be chosen from the results using detrended data and median values; however, we have seen that most of the methods for finding the seasonal patterns give similar results. One method might show a low in corn in September and another in October, but September and October will be the two lowest months. The seasonality, and risk, can also be confirmed by creating a seasonal volatility chart. All of these values can be seen in heating oil, Table 10.7, using the percentage monthly price changes. The median values are lower than the averages for May through September and December and January, indicating that there are many months in which extreme prices distort the averages.

TABLE 10.7 Monthly Heating Oil Prices as a Percentage Difference from the Previous Month, 1984–2011

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1984					-0.033	-0.046	-0.006	0.013	0.052	-0.080	0.000	-0.037
1985	-0.001	0.060	0.101	-0.080	-0.079	-0.022	0.000	0.086	0.073	0.069	0.022	-0.087
1986	-0.265	-0.061	-0.170	-0.011	-0.048	-0.140	-0.093	0.221	-0.030	-0.043	0.084	0.120
1987	0.091	-0.137	0.097	0.010	0.067	0.031	0.014	-0.064	0.054	0.068	-0.017	-0.084
1988	-0.049	-0.060	0.064	0.102	-0.074	-0.168	0.066	-0.034	-0.090	0.117	0.102	0.101
1989	0.000	0.010	0.046	-0.047	-0.051	-0.029	-0.022	0.091	0.115	-0.020	0.055	0.762
1990	-0.431	-0.041	0.013	-0.036	-0.095	-0.049	0.154	0.316	0.392	-0.125	0.031	-0.144
1991	-0.101	-0.050	-0.153	-0.043	0.028	-0.043	0.083	0.067	0.006	0.069	-0.102	-0.240
1992	0.140	-0.036	0.052	0.071	0.048	0.032	-0.022	-0.017	0.057	-0.060	-0.043	0.001
1993	-0.001	0.047	-0.059	0.001	-0.024	-0.036	-0.040	0.064	0.032	-0.058	-0.110	-0.024
1994	0.183	-0.002	-0.083	-0.025	0.018	0.027	0.031	-0.050	0.001	0.004	0.009	0.024
1995	-0.068	-0.005	0.007	0.071	-0.048	-0.030	0.020	0.032	0.002	0.023	0.036	0.150
1996	-0.143	0.210	0.062	-0.018	-0.171	-0.008	0.034	0.149	0.124	-0.045	0.053	-0.021
1997	0.006	-0.206	-0.003	0.076	-0.072	-0.049	0.046	-0.072	0.126	-0.007	-0.078	-0.071
1998	-0.005	-0.125	0.032	-0.034	-0.101	-0.001	-0.088	-0.003	0.186	-0.042	-0.180	0.003
1999	-0.008	0.010	0.354	0.004	-0.106	0.143	0.148	0.086	0.086	-0.073	0.168	0.053
2000	0.581	-0.270	-0.093	0.102	-0.059	0.131	-0.115	0.293	-0.044	0.004	0.085	-0.133
2001	-0.112	-0.063	0.056	-0.035	0.021	-0.100	0.004	0.125	-0.179	-0.050	-0.170	0.105
2002	-0.103	0.090	0.206	0.038	-0.100	0.080	0.019	0.088	0.067	-0.106	0.067	0.139
2003	0.141	0.206	-0.294	-0.056	-0.056	0.019	0.046	0.009	-0.060	0.021	0.079	0.106
2004	0.027	-0.002	-0.038	0.039	0.049	-0.026	0.185	-0.021	0.234	0.084	-0.039	-0.151
2005	0.148	0.042	0.105	-0.074	-0.030	0.103	0.021	0.261	0.025	-0.142	-0.120	0.070
2006	0.055	-0.062	0.121	0.051	0.002	-0.020	-0.010	-0.009	-0.122	-0.061	0.128	-0.094
2007	0.019	0.086	0.054	0.021	-0.018	0.140	-0.038	-0.020	0.109	0.072	0.071	0.022
2008	-0.033	0.117	0.107	0.031	0.136	0.058	-0.102	-0.095	-0.136	-0.277	-0.125	-0.259
2009	0.113	-0.094	0.032	-0.014	0.208	0.096	0.004	0.054	-0.087	0.213	-0.041	0.075
2010	-0.092	0.031	0.075	0.055	-0.108	0.005	0.006	0.000	0.082	0.024	0.051	0.056
2011	0.084	0.089										
% Up mos	50.0%	46.2%	72.0%	56.0%	34.6%	46.2%	65.4%	65.4%	73.1%	42.3%	61.5%	57.7%
Average	2.4%	0.9%	5.2%	2.8%	-1.3%	2.0%	3.6%	7.9%	6.5%	-1.6%	2.3%	3.6%
Median	-0.1%	-0.2%	4.9%	0.3%	-4.8%	-0.8%	0.6%	3.2%	5.2%	-2.0%	2.2%	0.3%
Max	58.1%	46.2%	72.0%	56.0%	34.6%	46.2%	65.4%	65.4%	73.1%	21.3%	61.5%	76.2%
Min	-43.1%	-27.0%	-29.4%	-8.0%	-17.1%	-16.8%	-11.5%	-9.5%	-17.9%	-27.7%	-18.0%	-25.9%

457
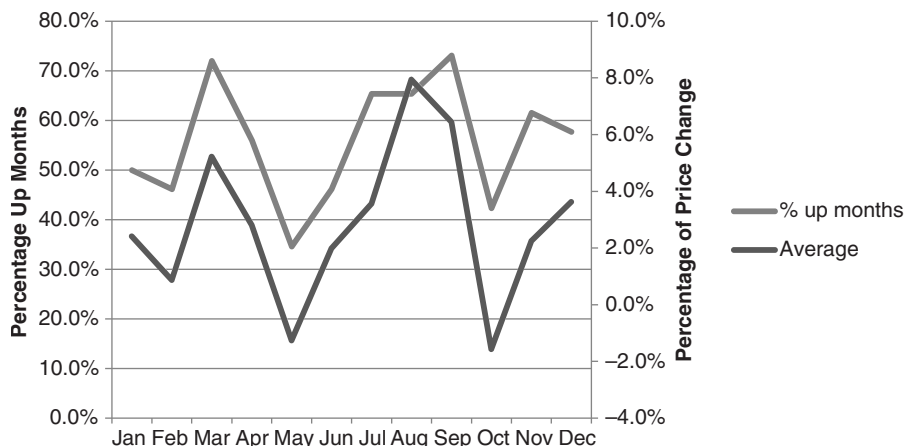
FIGURE 10.17   The percentage of upward months compared to the average monthly changes for heating oil, 1984 to 2011.

The results at the bottom of Table 10.7 can be visualized in Figure 10.17 where both the percentage of up months and the average monthly changes are plotted. Using only these numbers, the most basic seasonal trades would be

> *Hold longs* from the end of February to the end of March, from the end of May to the end of August, and from the end of October to the end of December.
> *Hold shorts* from the end of March to the end of May and from the end of August to the end of October.

Because these charts were based on average monthly cash prices, positions from the *end of February* should actually be averaged in over all of February, for example, 25% of the position each week. Using the price on the last day of the month, rather than the average, would make the results more practical for trading.

Of the three long positions, May through August represents inventory accumulation and October through December the critical winter months where users stock up on oil for the balance of the winter. Those trades seem reasonable. The first one, February and March may be statistically correct but still very risky. A late winter cold spell would drive prices sharply higher. Prices jumped 40% from March to April 1982, and 20% during the same period in 1983. We must also consider that the Gulf War, which pushed prices higher from August 1990, conformed by chance to the seasonal pattern; in reality, it could have happened any time during the year. Some traders may choose to trade the numbers only, but an understanding of the fundamentals might encourage further analysis and a better way of assessing whether the trade has a good reward to risk ratio.

Soybeans have similar issues. The classic seasonal trade is long from April 1 to June 1 (planting and crop development) and short from September 1 to November 1 (harvest). One commodity firm shows the highs in May/June and the lows in October and evaluates the seasonal moves as *a post-harvest rally into the new year, then a tax-related producer selling and the maturing of the Brazilian crop leading into the notorious*

February break. Further, the spring rally from March 1 often continues into the May/June planting, and the post-harvest rally begins by mid-November. This advice does not conflict with the standard seasonality.

Years with Similar Characteristics

Seasonal studies often yield results that are not as clear as desirable, and these results may be rejected because of the obvious lack of consistency. Often, this is caused by a few years that conflict with the normal seasonal patterns due to special events. If the month-end returns for cash corn are averaged from 1990 through 2010, the pattern is not quite as expected, as seen in Figure 10.18. Instead of the lows occurring only during harvest, in October, there is a low at the end of June as well as the end of September. A small rally in August reflects those years when there was uncertainty during the growing season; however, the primary pattern is that prices in most years just decline into harvest then rally into winter to reflect the carrying charges of storing the new crop combined with some export news.

These patterns can be more intuitive if the data is separated into *similar years*. For example, crop production is primarily determined by weather. Poor weather will cause sharp rallies during the growing season while good weather results in a dull, sideways market. The reaction to bad weather develops slowly. A drought is not caused by the first hot day but by prolonged days of sunshine and no rain. Similarly, delayed planting due to wet fields or a late winter will set the stage for an underdeveloped crop. A trader can see the characteristics of a developing weather market and enter a long position ahead of the rally.

Another important distinction is between bull and bear years. Cash corn prices will be used from 1970 through 2010. These two patterns can be separated a number of ways, but for this example the bull years are those where the average price of July through December is greater than the average price of January through June. Of those 41 years, 16 were bullish and 25 bearish. The average price for each month is shown in Figure 10.19. The bearish years have a "normal" seasonal pattern, while the bullish years
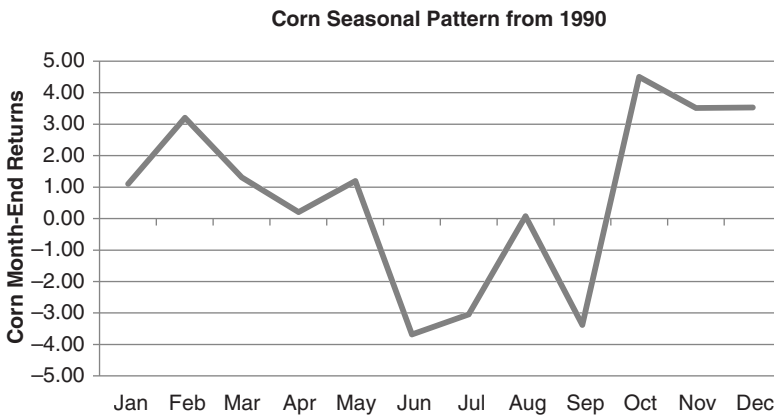
FIGURE 10.18 Month-end average returns for cash corn, 1990 to 2010.
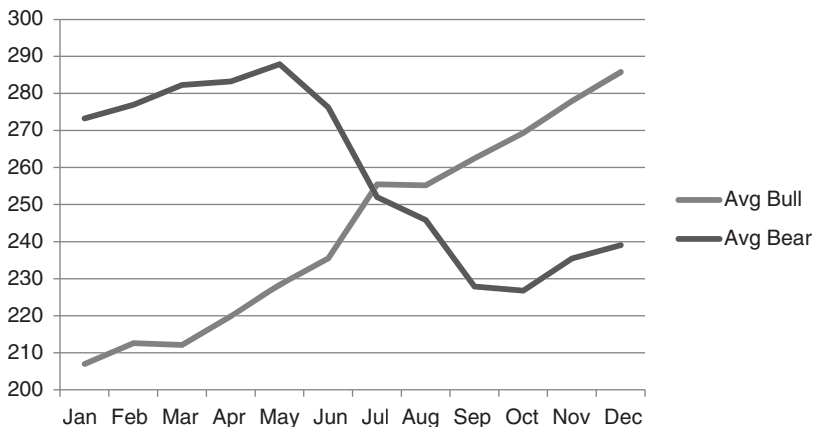
FIGURE 10.19 Corn monthly average prices separated into bull and bear years.

are consistently higher each month. Some of this uniformity can be attributed to larger upward moves in the most recent years.

Bullish and bearish years could be separated in other ways, the simplest being that the end of the calendar year is higher than the start of the year, or that the end of the crop year is higher than the start of the crop year. If we believe that normal seasonality causes low prices before planting, higher prices during the growing season, and lower prices at harvest, we can test the price this year against prior years to decide which pattern is now occurring. The current year prices can be compared against the previous 3-year or 5-year rolling average prices to decide relative value. This last method will be used in the next section.

Seasonal and Nonseasonal Market Patterns

There are years when agricultural prices follow a normal seasonal pattern and years that are quite different, just as there are bullish and bearish years. If a strategy is to profit from the rally during the growing season or the decline into harvest, then it's important to recognize whether the year is expected to have a normal pattern. There are many events that would cause nonseasonal patterns, such as large export commitments (increase in demand) or a previously bad harvest (decrease in supply), or even a large change in the value of the local currency. A poor harvest in the previous year will cause the new crop to start trading at a high price, in anticipation of continued short supply until the new crop yield can be assessed.

Using corn as an example because it represents the largest and most seasonal of the U.S. crops, we can construct a simple model that anticipates a seasonal or nonseasonal year and profits from those moves. It can be done by comparing the current cash prices with the highest and lowest prices of the past 5 years. A nonseasonal year would be one that has abnormally high prices in January. A seasonal year is one with relatively low prices in April, or high prices in July. The rules would be

For a seasonal year

If this year's price at the end of April is in the lower 50% of the high-low price range of the previous 5 years, then expect a normal seasonal pattern.

Buy on the last day of April, then *exit* on the last day of July.

If this year's price at the end of July is in the upper 35% of the high-low range of the past 5 years, then

Sell on the last day of July and *exit* on the last day of October.

For a nonseasonal year

If this year's price at the end of January is in the upper 35% of the high-low range for the past 5 years, then

Sell on the last day of January and *exit* on the last day of October.

Table 10.8 shows parts of the spreadsheet available on the Companion Website, *TSM Corn bull and bear years, seasonal patterns*. The full spreadsheet includes 1974–2010 month-end prices. In order to generate seasonal trades, it is necessary to create a series of seasonal calculations as shown in the tables. The first block has the month-end prices. The second is the 5-year high. Because the data start in 1974, the first year of the 5-year high is 1978. The third block is the 5-year low, and the last block is the position of this year's month-end price in the previous 5-year high-low range. That is calculated as, for example, similar to a raw stochastic,

100 × (January price – January 5-year low)/(January 5-year high – January 5-year low).

Once we have the relative position, the seasonal rules can be followed.

Table 10.9 shows the positions entered for the nonseasonal (bearish) trade, or the first seasonal (bullish) trade. In 2005 the price at the end of April was at the 8.1% level (out of 100%), near the low of the 5-year range; therefore, a seasonal long position was entered on the last day of April at 192.50. Seasonal trades entered in April are always closed out at the end of July, and, in 2005, that was at 215.50 for a profit of 23 cents, less costs. Because the July price was in the lower part of the previous 5-year range, 44.7%, no short was entered. Had it been above 50% a short would have been set at the July end-of-month price. Although there was a seasonal low in September–October, the relative low price in July made the second part of the seasonal trade too risky.

The seasonal trade in 2007 showed the opposite situation. Prices for the entire year were at 5-year highs; therefore no seasonal longs would be set. However, the end of July was also at the 100% level triggering a seasonal short trade on the last day of July, at 294. Unfortunately, prices continued higher in a nonseasonal pattern, generating a loss for the trade. Perhaps the 100% level indicates that risk is too high.

The nonseasonal trades look to profit from high prices at the beginning of the year and progressively lower prices during the year, reaching an expected seasonal low in October, following the new crop. Whatever short supply or strong demand caused prices to be high

TABLE 10.8 Spreadsheet Calculations for Seasonal and Nonseasonal Patterns

Month-End Prices

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2005	183.50	196.00	196.00	192.50	201.00	196.00	215.50	185.50	167.50	169.50	176.00	200.00
2006	203.00	211.50	215.50	221.50	222.50	215.50	213.00	214.50	237.50	305.00	367.00	365.00
2007	380.00	402.50	340.50	338.50	378.50	315.00	294.00	299.50	333.00	359.00	372.00	425.50
2008	469.50	519.00	532.00	566.50	555.50	680.50	537.50	534.50	441.00	364.50	325.00	378.50
2009	353.00	340.50	385.00	381.50	416.00	338.00	334.50	318.00	327.50	343.50	374.00	376.00
2010	328.50	358.50	321.50	350.50	334.50	332.00	365.00	393.50	470.50	561.00	515.50	590.50

5-Year High

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2005	263.50	287.50	302.50	306.50	293.50	248.50	236.50	253.00	238.00	239.50	235.50	235.50
2006	263.50	287.50	302.50	306.50	293.50	248.50	236.50	253.00	238.00	305.00	367.00	365.00
2007	380.00	402.50	340.50	338.50	378.50	315.00	294.00	299.50	333.00	359.00	372.00	425.50
2008	469.50	519.00	532.00	566.50	555.50	680.50	537.50	534.50	441.00	364.50	372.00	425.50
2009	469.50	519.00	532.00	566.50	555.50	680.50	537.50	534.50	441.00	364.50	374.00	425.50
2010	469.50	519.00	532.00	566.50	555.50	680.50	537.50	534.50	470.50	561.00	515.50	590.50

5-Year Low

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2005	183.50	191.00	183.00	182.50	175.00	174.00	198.50	185.50	167.50	169.50	176.00	188.50
2006	183.50	191.00	190.50	185.50	201.00	196.00	202.00	185.50	167.50	169.50	176.00	188.50
2007	183.50	196.00	196.00	192.50	201.00	196.00	202.00	185.50	167.50	169.50	176.00	188.50
2008	183.50	196.00	196.00	192.50	201.00	196.00	205.00	185.50	167.50	169.50	176.00	188.50
2009	183.50	196.00	196.00	192.50	201.00	196.00	213.00	185.50	167.50	169.50	176.00	200.00
2010	203.00	211.50	215.50	221.50	222.50	215.50	213.00	214.50	237.50	305.00	325.00	365.00

5-Year Relative Position

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2005	0.0	5.2	10.9	8.1	21.9	29.5	44.7	0.0	0.0	0.0	0.0	24.5
2006	24.4	21.2	22.3	29.8	23.2	37.1	31.9	43.0	99.3	100.0	100.0	100.0
2007	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2008	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	76.0	80.2
2009	59.3	44.7	56.3	50.5	60.6	29.3	37.4	38.0	58.5	89.2	100.0	78.0
2010	47.1	47.8	33.5	37.4	33.6	25.1	46.8	55.9	100.0	100.0	100.0	100.0

TABLE 10.9 Positions, Entry and Exit Prices, Profits and Losses for Seasonal Patterns

Position		Seasonal Trade #1		Seasonal Trade #2		Profit/Loss		Total Profit/Loss			Total
		Entry	Exit	Reverse	Exit	Trade 1	Trade 2	Trade 1	Trade 2	1 + 2	All Patterns
2005	1	192.5	169.5	0	0	−23.0	0	−121.0	299.5	178.5	708.5
2006	1	221.5	305.0	0	0	83.5	0	−37.5	299.5	262.0	792.0
2007	−1	0	0	294.0	359.0	0	−65.0	−37.5	234.5	197.0	748.0
2008	−1	0	0	537.5	364.5	0	173.0	−37.5	407.5	370.0	1026.0
2009	0	0	0	0	0	0	0	−37.5	407.5	370.0	1026.0
2010	1	350.5	561.0	0	0	210.5	0	173.0	407.5	580.5	1236.5

TABLE 10.10 Positions, Entry and Exit Prices, and Profits and Losses for Nonseasonal Patterns

Year	Position	Entry	Exit	PL	PL Bear	Cum Bear
2005	1	192.5	169.5	–23.0	0	530
2006	1	221.5	305.0	83.5	0	530
2007	–1	380.0	359.0	21.0	21	551
2008	–1	469.5	364.5	105.0	105	656
2009	0	0.0	0.0	0.0	0	656
2010	1	350.5	561.0	210.5	0	656

in the previous year is likely to correct with the new crop. The first year in the table that a nonseasonal short is entered is 2007, shown in Table 10.10, where prices were at 5-year highs. Although they stayed at 5-year highs all year, prices did drop from the January level of 380 to the October harvest level of 369. It was not much of a correction, but still a net profit. Had we waited longer, the small profit would have disappeared completely.

Figure 10.20 shows the performance of this method over 37 years, separating the results into the first and second seasonal trade (long then short), and the nonseasonal trade (short), then combining the results into the total profit (the average). Performance is good although there are years when there were no trades, as shown by the horizontal lines. Remember that this was intended to be a simple example to show that, in theory, grains follow a reasonably predictable pattern, helped along by nature.

Not Quite So Easy

However, to achieve these profits you would need to buy and sell cash corn. That is, when you buy corn you would be holding a warehouse receipt for 5,000 bushels of corn for each contract lot size. Selling short in January or July means promising to deliver corn at some future date. For most traders, that is not practical and perhaps impossible. The only choice is trading futures, but the futures price is never the same as the cash
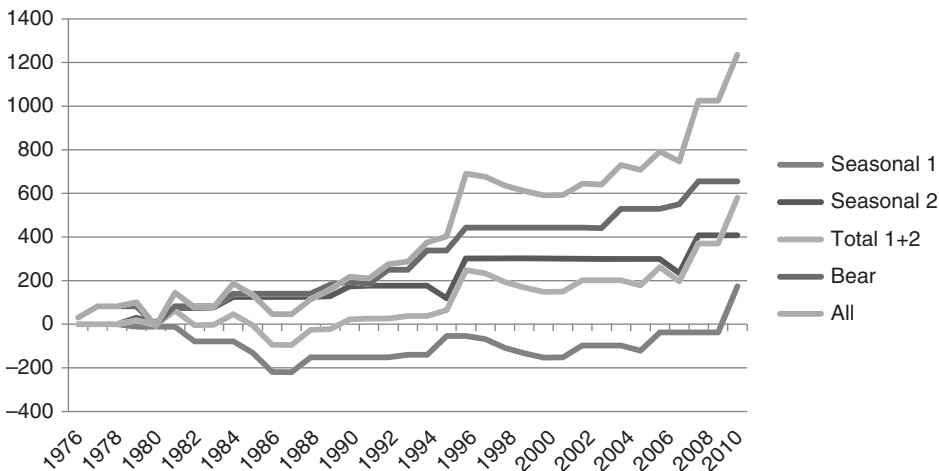
FIGURE 10.20 Performance of a seasonal system applied to cash corn prices.

price. It discounts future expectations. If everyone has the same opinion of what will happen between April and July, then on the last day of April, the price of December corn (the new crop) will already be much higher than the cash price, reflecting uncertainty during the growing year. Or, if a good new harvest was expected, then the December futures price would reflect the harvest lows as soon as possible.

A more realistic scenario can be found by using only the December futures contract, the first new crop month. Each year, the new crop prices start on December 1 and trade until the last day of November in the following year. A seasonal chart can be constructed by taking the month-end differences just as before, but the November–December differences can't be used because they reflect different contracts. A spreadsheet with this data and the seasonal chart can be found in the Companion Website along with all of the other seasonal charts in this chapter.

Figure 10.21 shows the month-end average changes, which are quite different from the cash pattern. The lows are in June and September, indicating that traders normally anticipate a good crop and discount the final price as early as June before being surprised that the summer is still uncertain. This pattern can now be traded because it uses only December futures. Taking the obvious two major trends, the following seasonal trades can be isolated:

1. *Sell* December corn at the end of February, *exit* at the end of July.
2. *Buy* December corn at the end of July, *exit* at the end of October.

Neither of these trades span the November–December months when we would roll into the next year's new crop. The results of using these two seasonal trades are simple to calculate and are shown in Figure 10.22 and in the Companion Website as *TSM Dec Corn Strategy*. Surprisingly, the first seasonal trade, which expects prices to decline from February through July, is consistently profitable. The weather scares that always surround planting (too much rain), and then growing season (not enough rain) do not seem
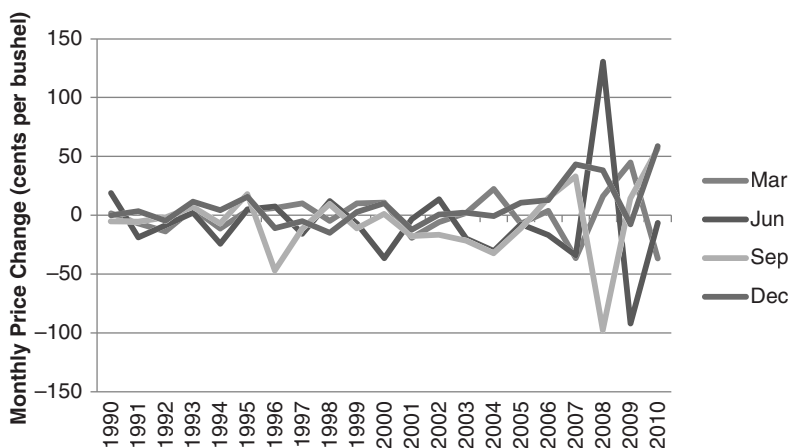
FIGURE 10.21   December corn contract, monthly average changes from 1990 to 2010.
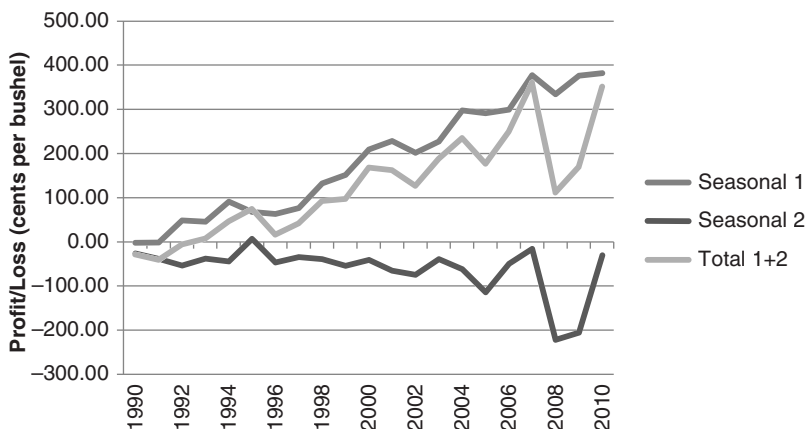
FIGURE 10.22 Results of two seasonal strategies using December corn futures.

to be significant. While there are some years with small losses, the overall performance of Strategy 1 is good. Not so with Strategy 2. Although it is perfectly logical to expect a price rally after harvest based on both storage costs (carrying charges) and export activity, the results of that strategy is at best a break-even with more volatility.

Effects of Currency (Argentina)

There was a time when the United States was the biggest exporter of grain and when currencies did not fluctuate very much. But at that time many currencies were backed by gold and not freely traded. Argentina is now the biggest producer in South America that has excess production; therefore, it competes directly with the United States on the world market. Let us assume the U.K. wants to import wheat and, for the sake of simplicity, we assume the value of the British pound remains the same against the U.S. dollar. Which country would the British buy from, the United States or Argentina?

Looking at the changing value of the Argentine peso in Figure 10.23, there is a variation of 120%, from 5 to 11 pesos per USD, over the one year from May 2009 to May 2010. At the low point, 5 pesos to the dollar, the peso was strong and the price of Argentine wheat high; therefore, U.S. wheat is more attractive. At the high point, 11 pesos to the dollar, the peso would have been cheap, and the price of wheat in Argentina would have been attractive. There are many years in which the change in currency value is greater than the change in the product value. In all cases, when you combine the price fluctuations with the currency fluctuations, you get move volatility.

Using Continuous Futures

Knowing that only futures (or options on futures) will be traded, we looked at the seasonal patterns based on only trading one contract each year, December, the first new crop for corn. For most analysts, continuous, back-adjusted prices are the most readily available and easiest data to use. They also combine the most liquid part of each delivery month.
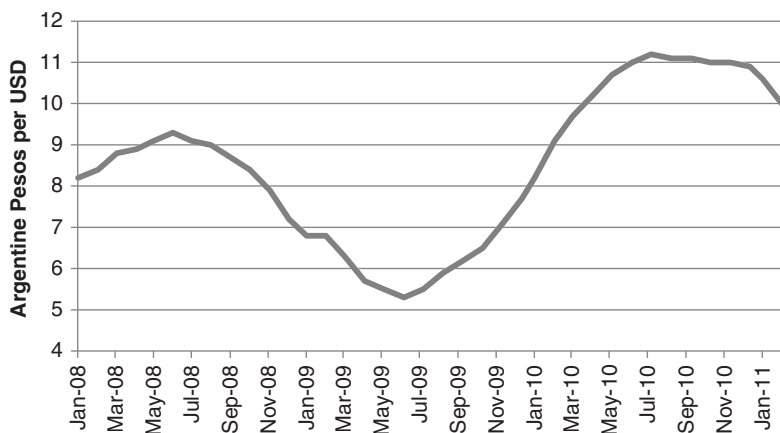
FIGURE 10.23   Argentine pesos per USD show large cyclic swings during the past 3 years.

In the interest rate markets, forex, metals, and equity index markets, back-adjusted data works for most strategies, and for trending methods in any market, back-adjusted data are fine. However, there are some strategies, such as arbitrage and seasonal trades, that need the exact prices at the time. And, for any percentage calculation that uses prices, a back-adjusted price gives the wrong result.

Figure 10.24 shows the back-adjusted corn series using futures beginning in 1949. Prices on the far right are 2011, and are the same prices trading at this time. Each earlier contract is adjusted up or down according to the price gap between the two contracts on the roll date. Prices that become higher in the past are the result of upwards gap adjustments. That is, the more recent contract is trading at a higher price than the older contract. That would be normal for all corn contracts because it is a carrying-charge market (in contango) except for the switch from old crop to new crop, usually in December. Then
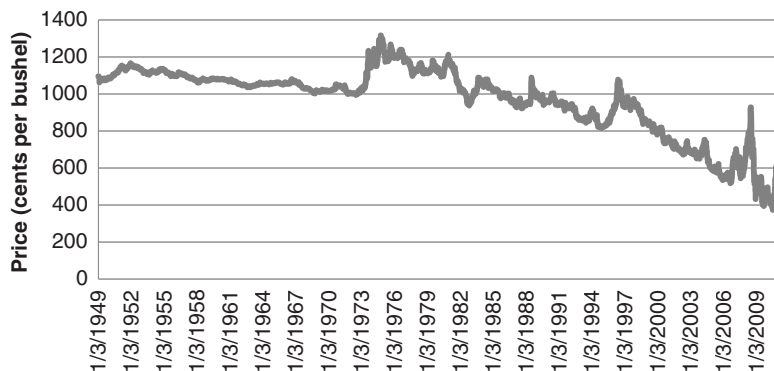
FIGURE 10.24   Back-adjusted corn prices suffer from upward gaps at most roll dates, causing historic prices to be much higher.

rolling into December from March, March from May, and July from September will usually require adding an adjustment factor to the older contract to raise those prices to the level of the newer data. The new crop, December, should be trading lower than the last old crop contract, September, so that adjustment is often downward. Given this very different historic picture of corn prices, can this data be used for trading seasonal patterns?

As before, a seasonal chart can be constructed using back-adjusted futures, but only by subtracting the price of the previous month from the current month. Unlike cash, we cannot find the percent difference in price change. Using price differences adds a significant bias to more recent data, which have been much more volatile than past data. Figure 10.25 shows the average monthly price change over the 21 years from 1990 as well as the averages of the first 11 and last 10 year periods. The pattern is similar to Figure 10.21, which used only December futures. The seasonal pattern shows a low at harvest, the peak prices are not in June or July, and the volatility of the recent 10 years is far greater than the earlier 11 years. This reduces the importance of older data, but many analysts prefer giving more weight to more recent data. As long as you understand the biases of the data, you can work with it.

The seasonal pattern based on futures suggests two clear strategies, the first is the same as the one using only December futures, but the second takes advantage of the month of December, which could not be used in the previous strategy because it was a different crop year.

Sell at the end of February and exit at the end of July.
Buy at the end of July and exit at the end of December.

If these two methods are implemented, although with the complete benefit of hindsight, the results are shown in Figure 10.26. Strategy 1 is consistently profitable, as it was using only the December futures, while strategy 2 is slightly better when the month of December is used.
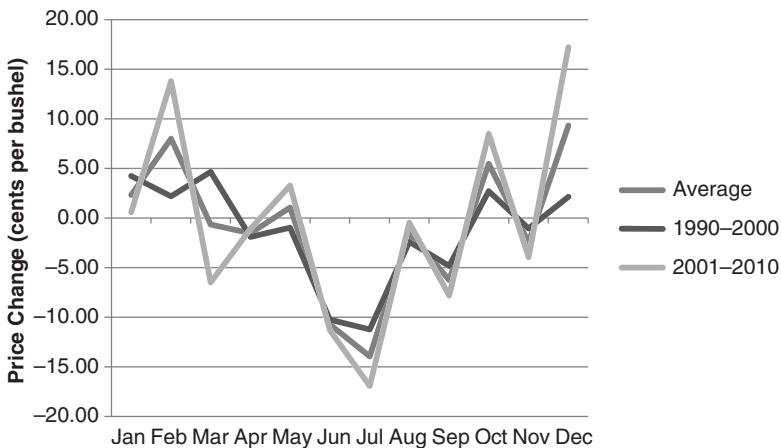
FIGURE 10.25 Corn monthly average price changes using futures market back-adjusted data.
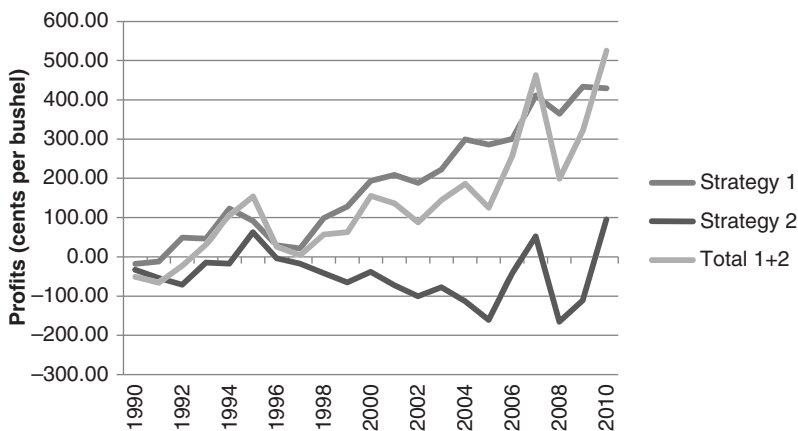
FIGURE 10.26   Performance of two seasonal strategies for corn based on back-adjusted futures prices.

Strategy 2, buying at the end of July, might be justified if the new crop size is known by then, which is possible. Then exiting at the end of December could capture the post-harvest rally. The results, however, show that this is inconsistent, even though the seasonal pattern in Figure 10.25 seemed clear.

The first method, selling at the end of January and exiting at the end of July seems completely contrary to seasonality but consistent with the idea of a nonseasonal year, where prices start high and end at harvests' low. Using cash prices, July was expected to be the high. The consistency of the results, with 12 of 21 years profitable but the profits much larger than the losses, seems to justify the method. The confusing part is "Why is July the seasonal low?"

An answer can be found by looking at the increasing volatility of monthly changes over the past 20 years, as seen in Figure 10.27. While volatility had been increasing at an
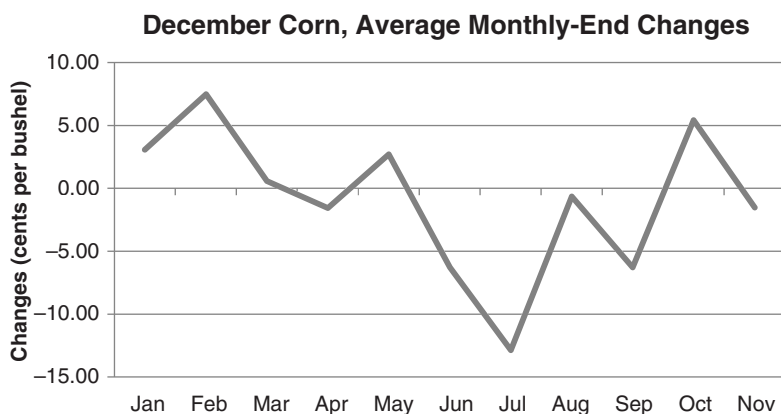
FIGURE 10.27   Increasing volatility in price changes for corn, based on the March, June, September, and December futures contracts. Volatility has been increasing but has been extreme during the past 3 years.

orderly rate until 2007, it has been extreme for the past few years. Using back-adjusted futures means that we cannot divide the monthly price change by the futures price; therefore, the changes in the most recent years will overwhelm the older years. One solution is to divide the monthly price changes by the unadjusted futures prices, but that requires constructing a continuous but not back-adjusted series. That can be done using a data service such as CSI.

Seasonal Patterns in Stocks

Carrying the same analysis into the equities markets, we can find many companies that are likely to have seasonal patterns. These will not be diversified holding companies, but those whose primary income is from a single, seasonal source. Many airlines qualify, as do firms in the travel and leisure sector. It is not clear that share prices of the major oil companies, such as Exxon-Mobil, vary directly with the price of oil. Refining margins and vertical integration of retail gas stations (with mini-markets) make the profitability of these companies much less dependent upon the varying cost of oil.

Using AMR, the parent company of American Airlines, as an example, the familiar monthly returns are shown in Table 10.11 from September 1998 through April 2011. Along the bottom are the monthly averages as well as a selected set of averages, This selected set removed September and October 2001 as well as September 2008 through March of 2009. The first months removed reflect the 9/11 terrorist attack, in which American Airlines played an unwanted role. The second set isolates the worst part of the sub-prime mortgage meltdown, where business came to a near halt. When we remove the outliers we can be guilty of the same data manipulation that caused the end of Long Term

TABLE 10.11 Monthly Returns of AMR with Monthly Averages and Averages That Removed Selected Months

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1998									1.73	20.96	−1.66	−9.96
1999	−1.07	−5.63	5.62	19.24	−6.80	4.90	−4.95	−9.65	−7.05	16.51	−4.13	10.09
2000	−19.70	−1.71	42.38	6.84	−16.32	−7.23	25.04	−0.76	−0.18	0.00	2.11	17.19
2001	−0.26	−14.94	5.62	8.51	2.31	−7.34	−2.71	−8.99	−40.17	−4.91	17.36	4.40
2002	11.84	4.65	1.19	−18.71	−2.42	−19.52	−33.69	−8.86	−58.98	12.92	64.19	−14.84
2003	−56.06	−19.31	−10.26	113.33	41.52	73.50	−15.00	17.65	4.09	15.98	−3.39	0.94
2004	26.64	−7.32	−16.25	−10.84	1.50	5.12	−30.39	6.05	−18.01	5.32	16.97	21.26
2005	−21.46	−1.28	26.03	−2.15	23.21	−6.12	16.02	−10.39	−11.20	20.84	25.02	31.62
2006	2.11	10.57	7.77	−8.91	0.08	3.08	−13.45	−6.14	12.06	22.47	12.77	−5.41
2007	22.56	−8.02	−10.65	−14.32	8.66	−7.05	−6.34	−0.69	−9.06	7.67	−11.75	−33.76
2008	−0.64	−8.11	−29.59	−2.77	−18.02	−28.79	76.37	14.40	−4.94	3.97	−14.01	21.53
2009	−44.33	−31.14	−22.00	49.22	−6.51	−9.66	33.08	2.06	45.60	−32.20	12.06	27.98
2010	−10.48	32.80	−0.87	−18.99	3.93	−11.60	4.42	−13.56	2.45	25.84	8.49	−9.00
2011	−9.50	−4.40	−4.15									
Average	−7.72	−4.14	−0.40	10.04	2.59	−0.89	4.03	−1.57	−6.43	8.88	9.54	4.77
Selected	−4.67	−1.89	1.40	10.04	2.59	−0.89	4.03	−1.57	−3.50	10.57	11.50	3.38
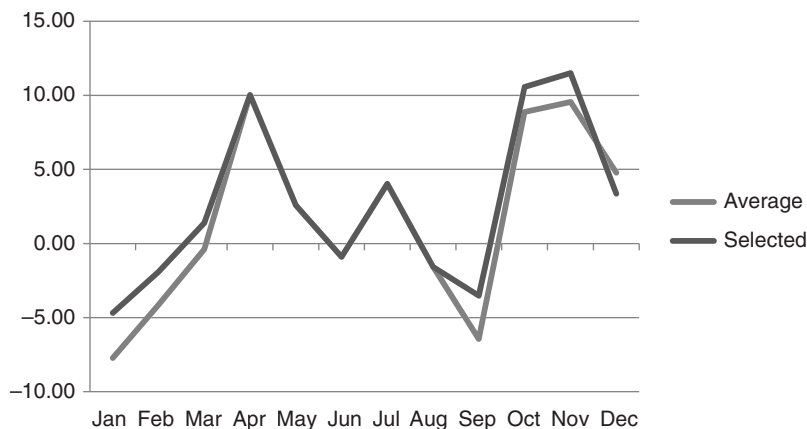
FIGURE 10.28   Seasonal patterns of AMR, all months and selected months of extreme crises removed.

Capital Management in the late 1990s. *Caveat emptor*. Both monthly patterns are shown in Figure 10.28.

Surprisingly, the removal of critical months made little difference to the pattern, even over 13 years, which is a small period to judge seasonality. There seems to be a clear pattern of prices peaking in April and November–December, and falling to lows in September. We can guess that the year-end holiday season is the cause of strength, but we don't know if the other patterns reflect the time when customers were actually flying or if it is when they purchased their tickets. We expect the summer to be a busy time for airlines; however those purchases may have come in April.

Seasonality for Trading

Before analyzing a market for seasonal patterns, you must understand whether that market qualifies. Given the power of computers and the convenience of spreadsheets, it is possible to evaluate all stocks and all futures markets looking for patterns. Finding a monthly pattern does not mean that there is seasonality. Years ago, when computing power was just taking hold, there was an academic who discovered that, for the past 18 years, whenever there was a good coffee crop in South America, then next year there was a good wheat harvest in Europe. He delved deeper and found that the Gulf Stream, a well-known flow of warm water, passes by the eastern coast of South American and flows north all the way to Europe. Voilà! Having decided that the chances of 18 consecutive years of success was convincing, combined with a fundamental justification, he sold wheat the next year following a good coffee crop. Of course, that was the first year that the theory failed. It turns out that he had tested hundreds of combinations of causal effects, a crop in one location followed by a crop in another location, in order to find 18 consecutive years. We understand now that it was an early case of overfitting; however,

the lesson to be learned is that there must be a clear fundamental reason to believe that there is a seasonal pattern. The computer is used to validate a belief, not to create one.

By way of example, there is an interesting pattern that develops in livestock during a period of high grain prices. Most often, a farmer who also feeds cattle will send them to market early rather than pay the cost of high feed and then face the uncertainty of a profit when the cattle are ready to market. The decision to sell early puts immediate pressure on cattle prices in the short term but causes higher prices 3 to 9 months out, when there will be a shortage of supply. As grain prices move higher, selling the nearby cattle futures and buying the deferred can be a profitable strategy.

Although metals are not subject to seasonal supply factors, copper, as well as most of the LME nonferrous metals, should have a significant seasonal demand component. All of them are primarily used in housing for plumbing, sheet metal, or stainless steel. More recently, commodities in general have seen an unprecedented rally due to fears of inflation as well as demand from the rising middle class in China, with India soon to follow. Seasonality must be hidden within these prices; however, the larger macroeconomic factors have overwhelmed them to the point where they are insignificant for now. Just as in the comparison of corn cash prices and futures prices, it is necessary to be realistic about which patterns to use.

The systems and methods that follow are based entirely on seasonal patterns and may be used alone or as a filter for other strategies.

Seasonal Studies and Key Dates

Most agricultural commodities exhibit traditional, reliable price moves at one or more periods throughout the year. The grains grown in the northern hemisphere have a high likelihood of a rally during the late spring and early summer, when the chance of drought has the greatest effect on yield. When prices show the normal harvest lows, followed by a modest rise and sideways pattern throughout the winter, the potential is good for a rally during the early growing season. When prices begin the new crop year at relatively low levels, the risk of holding a long position is small. Once prices have moved higher, there is rarely a season where a short sale of corn, soybeans, cotton, or sugar will not net a good profit within the two months before harvest begins.

Seasonal studies are intended to provide information on when the largest move will occur. The following studies—Grushcow and Smith (1980), ContiCommodity (1983), and Bernstein (1986)—can be used to compare historic seasonality in some of the more active commodities. A summary of the results is shown in Table 10.12. Each study offers a different perspective on seasonality. Grushcow and Smith analyze both cash and individual futures markets over a fairly long period and present complete statistics; ContiCommodity used mostly the past 10 years (which ended in 1981), but included a unique volatility analysis; Bernstein, the most recently published study, gave the most complete background on calculations, including separate studies of bullish and bearish years and an exceptionally long time period for cash market analysis.

TABLE 10.12 Results of Seasonal Studies

	J-F	F-M	M-A	A-M	M-J	J-J	J-A	A-S	S-O	O-N	N-D	D-J
Corn												
24 years*	.5	−1.1	1.6	3.6	3.4	.5	1.5	−5.3	−8.5	−1.3	.5	2.1
10 years†	2.2	−1.3	−2.3	−1.2	2.3	7.7	4.5	1.4	−7.5	−1.3	−6.4	1.8
46 years‡	−.6	2.0	3.0	2.9	2.3	3.8	−1.5	−3.2	−6.7	−3.7	5.4	1.5
%*	58	58	67	92	50	46	50	54	88	63	83	71
%‡	52	60	80	72	45	59	62	63	82	60	82	70
Soybeans												
24 years*	7.4	7.5	8.4	9.5	6.4	−.7	−1.3	−25.8	−7.5	4.2	5.8	.3
10 years†	−2.3	21.1	23.7	12.6	20.0	14.7	−11.7	−14.9	−46.7	−9.8	−8.9	2.1
42 years‡§	3	11	8	8	3	−1	−5	−20	−8	10	3	2
%*	67	50	67	50	42	38	58	83	63	75	67	79
%‡	62	53	57	51	42	37	64	74	64	76	72	70
Cattle												
25 years*	−.4	.1	.6	.1	−.1	.4	.4	−.3	−.4	−.5	−.1	.4
10 years†	1.4	.2	.8	1.8	.4	−1.2	−.4	−.7	−1.3	−.5	−.3	−.2
50 years‡§	1	12	13	4	−1	7	3	4	−9	−9	−2	2
%*	64	60	48	44	56	60	56	44	76	64	52	72
%‡	52	70	59	52	52	59	55	42	60	71	51	62
Orange Juice												
10 years*	−1.6	−.1	−.3	−.9	−.9	−.3	.1	−.2	1.2	−2.3	−.3	1.7
10 years†	−1.9	2.4	−1.1	−.7	.0	1.6	1.8	2.3	.2	.0	.0	−4.7
34 years‡	27	14	4	−3	−8	−1	1	3	−2	−2	−13	−4
%*	78	45	67	56	45	45	56	67	45	45	56	56
%‡	6!	64	50	58	50	65	66	81	61	40	67	62
Coffee												
22 years*	−.2	.9	−.5	−.4	.1	.1	−.7	−.3	.1	.3	−.5	.1
8 years†	−6.5	1.8	7.6	5.0	−.3	3.3	−11.5	−2.0	4.3	−8.6	4.7	2.2
53 years‡§	0	−10	2	−4	1	4	18	2	19	−10	−4	−2
%*	41	64	50	50	59	46	46	55	41	64	50	77
%‡	56	58	43	53	50	48	56	56	51	59	54	49

*Grushcow and Smith (change in price), data ending 1978.
†ContiCommodity (% change in seasonal factor), data ending 1981.
‡Bernstein (% change in seasonal factor), data ending 1985.
§Approximate values.
% refers to the reliability of monthly seasonality.

The number of years in the seasonal analysis counts heavily in determining the normal patterns. As Table 10.12 shows, the ContiCommodity results, based on only 10 years, are often quite different from the other two studies. For trading safety, it would be best to select those patterns that have proved reliable over many years; however, because the most recent 10 years may be most important for deciding if the pattern is still profitable, a trader must be able to identify nonseasonal, bullish, or bearish patterns.

The conclusions that seem consistent throughout all studies are:

Corn and soybeans. September and October show major harvest pressure.
Cattle. An end of the year liquidation and mid-winter rally.
Coffee and juice. No common moves in the three studies.

Some commodities are more seasonally consistent than others. Both the coffee and orange juice markets were expected to show patterns that reflect a rise as the possibilities of a freeze increase; however, those patterns did not appear. Because we know that the freeze concerns must appear in the prices, we can infer that the normal seasonality of these products is distorted by the inconsistent and highly volatile periods that follow a freeze. These markets would be candidates for the bullish and bearish years study in which you compare only those years with common factors.

The three studies shown here, as well as most others, include recommended trades based on key dates, which reflect the patterns in Table 10.12. By selecting those trades common to all of them, you have found those that are most reliable. In a summary by Bernstein, which catalogs commodities by those months with the highest reliability, the agricultural products are clearly the most seasonal. The only nonagricultural market that shows any consistent seasonality is copper. Although there may be interesting arguments for the forces of demand on silver, currencies, and financial markets, their inconsistency shows that they are not candidates for seasonal trading.

Seasonal Calendar

Because Bernstein's work covers the cash markets over an extremely long period, it must be considered the most reliable source of basic seasonal patterns. Table 10.13 is part of the weekly seasonal calendar that appears in *Seasonal Concepts in Futures Trading.* The numbers in the table show those weeks with consistent historic moves. Weeks of 64% and higher represent upwards moves; weeks of 36% and lower are downward trends. This calendar can be extremely useful when combined with some simple trading logic that asks, "Is the market acting in a seasonal manner?" *before* the position is entered. It should also be noted that patterns in futures markets, whether a single contract or back-adjusted data, will be different from the cash market. If you are trading futures, then you need to use futures patterns.

Books on Seasonal Patterns

It is not possible to summarize all the literature on seasonality here; however, there are a number of books that can be studied to take advantage of work already done.

> *Seasonal Stock Market Trends: The Definitive Guide to Calendar-Based Stock Market Trading* by Jay Kaeppel (Hoboken, NJ: John Wiley & Sons, 2009).
> *The Almanac Investor: Profit from Market History and Seasonal Trends* by Jeffrey A. Hirsch and J. Taylor Brown (Stock Traders Almanac, 2005).

TABLE 10.13 Seasonal Calendar

	Mar Corn	May Corn	Jul Corn	Sep Corn	Dec Corn	Mar Wheat	May Wheat	Jul Wheat	Sep Wheat	Dec Wheat	Mar Oats	May Oats	Jul Oats	Sep Oats	Dec Oats	Jan Soybeans	Mar Soybeans	May Soybeans
1 Jan				76			29				66	66					29	
2												66		66				
3	66							22	23	37							66	
4							27			33	37						37	
5 Feb						31		35	37	31								66
6		27				64				75	66							
7	35										31	37	33					
8	35	29				22	27				27						70	
9 Mar	66						35		31	25			23	37	37		35	
10								27	37	31								
11			72			35		35	23	29								
12																		37
13 Apr			66	66			72		79			76	64	81				78
14		77										66	64					70
15							37		33									35
16							33							25	37			26
17			27					22				22						
18 May							37		29	18		83						76
19	35											82	76					
20	70				66		33											
21						37	33										66	
22 Jun				37				31	29					37		64	35	
23	33		33	35		76		35	28				37	37				
24	76		76		33			64	82								70	
25	33			27	36			37	29							76		64
26 Jul	66		70	66	63								64					
27			66					64					29	26				
28		35						64						35		13		23
29	22	22			21									31		25	17	29
30		27	23		36		27											
31 Aug				35	33													
32												70		66				
33												66		66				
34							83		64				64	66	66	66	37	
35 Sep				35			72		70	64		66		37				
36									64									66
37			31									70					29	29
38	29				31										35		33	
39																		
40 Oct												64			64	76		64
41						66		72		64							25	35
42				63								66						
43							27	33			70							
44 Nov			27	66	37	33		26				72				31	33	
45									35						79	64		
46																29	33	35
47			64					64										
48 Dec							29	35							66	64		
49	18		35								33				66			
50												35				29	66	25
51												64		35	35	27	37	
52							66											

TABLE 10.13 *(continued)*

Jul Soybeans	Aug Soybeans	Sep Soybeans	Nov Soybeans	Jan Soybean Meal	Mar Soybean Meal	May Soybean Meal	Jul Soybean Meal	Sep Soybean Meal	Dec Soybean Meal	Mar Soybean Oil	May Soybean Oil	Jul Soybean Oil	Sep Soybean Oil	Dec Soybean Oil	Feb Live Cattle	Apr Live Cattle	Jun Live Cattle	Aug Live Cattle	Oct Live Cattle
37							33			35									
64	64									66						64			
								29			77	66					64		
	37					33			66	66					66			66	
	64	70			28		37		37	70	64	62			64		76		
					37										66			75	
64	64												66		76		64		
						35				33								37	
	35				66	35	23	23	35						77	76	66		
					66	64		70		70	66							66	66
					37	35									66	70			
37										33						37			
	64					70	70			72					72	76	71		
64		63						70	66						64	66	76		
	25								37								64	75	66
23						29	17			33		66							
			36					29	29	70					64			66	
75										27							64		33
70	64	66				64			70						70				66
								64	75	27		33					70	75	66
23	33	37		66		66	31		31			27	33	22		64		64	
64									76				33	66	66				
37			64	81	66			66	66	70	33		66	66					33
	70		72								72							66	
75			66			64			66	70		72		72	66	77		73	
	64		66		76				64							37			77
		63			66	29			64						33				
	33	25	15	33	20	35		29	25		29								
	28							35	35				64						
								35	35									66	
			37						37										
								70	75	33									
	37				37			29	35	27	64		29	10					37
64	35					64					64			64	66	35			70
29		66				29		76	66	27					37				
									37						66				
																37			
64			36	72										35	33				
35			73		66		36		76	64					72				
		35								23							27	31	
									64						33	35			
64			77	66		70									70	72	13	64	
	76		75			64										64	70	75	
				66							16				66	70	76		
	37					35		37	29		66		27					66	
			66		66	72		66	33				33						
										70				33					
33	29			35	31					66	29				66	66			
37	35	31								11	27	22			64			66	

475

TABLE 10.13 *(continued)*

	Dec Live Cattle	Mar Feeder Cattle	Apr Feeder Cattle	Sep Feeder Cattle	Oct Feeder Cattle	Nov Feeder Cattle	Feb Live Hogs	Apr Live Hogs	Jun Live Hogs	Jul Live Hogs	Aug Live Hogs	Oct Live Hogs	Dec Live Hogs	Feb Pork Bellies	Mar Pork Bellies	May Pork Bellies	Jul Pork Bellies	Aug Pork Bellies	Mar Cocoa	May Cocoa	Jul Cocoa	Dec Cocoa
1 Jan									66					64		37						
2							70	77			76			70	66							
3			66						66	80	76				66	64	64		35		33	
4									73	64							66	62				
5 Feb		70								64	66			64	81							
6	66						64								37				64			
7																	35		35	33		
8		20		36				35	66									66		66		
9 Mar	80	80	70	63				66									70		75		66	
10				63	70																	
11				35					33										64	35	27	27
12	37			63								73					66				68	
13 Apr									66		70	64						79				
14	72		90	80	63	90		77	78	87		71			33		31				38	
15			80	81	72	66						33					66	78		66		
16				81													23	23		22	38	
17										64	64											
18 May				30	33	30			66	73												
19					27				80								27		37	29	27	
20				70						29	31						35			64	66	
21				63	81	80											17	35				33
22 Jun						77					35	26	27	35			35	23				
23	33			36		30				31					37				37			33
24	70			36						70				70			70				66	
25			72	63							64	66					70			76	77	
26 Jul	66			72	66						35				66		75					72
27				36	80									64					64	77	66	72
28				36	27												66				77	64
29											35				37							
30				27																		
31 Aug				36								66					64		64	66		
32				72	63						76	66	72						64		36	
33				36								70			76	66		64		33		
34				63		30						37	38						64	83		
35 Sep	70			63	80							73	66	70	66		81		70			77
36				63	72	70							66							33		66
37						70									37							
38						30																
39	27																			33		
40 Oct	33				70							31		29		33	29	23				
41													27		31					29	33	
42	35			36													37		70			
43	22													66	33		75			72	64	
44 Nov							82	83	71					72	64	66	72		66	70	29	
45	70													77	70		80		25	33	33	
46		70					84						77									27
47																						
48 Dec							64							35			35		64	66		72
49	33		66										66									
50													76		37	66	64					
51									64				73			77	76					
52																						

476

TABLE 10.13 (continued)

Jan Orange Juice	May Orange Juice	Sep Orange Juice	Nov Orange Juice	Mar Sugar	May Sugar	Jul Sugar	Oct Sugar	Mar Coffee	May Coffee	Jul Coffee	Sep Coffee	Dec Coffee	Mar Cotton	May Cotton	Jul Cotton	Oct Cotton	Dec Cotton	Jan Lumber	Mar Lumber	May Lumber	Jul Lumber	Sep Lumber	Nov Lumber
31	35	29								63					75	87	86		66				
	35				77	70		63		66						75			66			28	
37								27	18	27					64				30		69		
	35			35						66	75					35	33				69	64	
		35				35			63	63	72	81							71	69		69	66
						70			66	91	75	63							15			64	63
		35						72		33		70	35	37	18	37							76
	35	29	33	62	72	35		36	25		66		64	64	70	68			21	7	28	72	28
									70	66		72							30			35	16
25		31		64	29	37	36	36	33		33				37	35	31					33	
26	29	12	16		27	37	36	80	75		83	72			35				35	69			30
				22						66	33				70		75			15		25	35
31	64		70							75	36	36			37							36	7
31		37	37					36		66	66	66	72		66					70		33	
75					33	37					63	63								70			
75	64	76	76					63	66	75		27			29	26							
																		30					
		77			33			36	72		72	80		75		68			27	30	7	25	23
						35			72	66	63	76			70						35	33	35
68		35						36		75	72	72			70	66			23			25	
68										66	36	35			37	31	35	69			28	83	14
33		64				64				29						29					78	33	83
35				33	33			36	36		33				70	37		66					35
		35						33	36		27	70				75	69					75	64
66		35	27		29			36			33												
					35							36	64										
73		66						36				68				37		37	30				35
	64				29							70		35		72			15		35		28
		64						36	63		66				76				?			66	71
								36					64										
	70									63	66	80											28
37								72		66	63	37				37							71
73			66					66	27		36	64	64			37						66	
73		82	66	76				36	75			36				66							
		70	66					72		66	36			35	35					30			69
						36				63	33	63	35	35		37			25	7		25	28
75					27	37	36	63	66				63	23	37	29	12			35	23	28	
								63					23										
81	82	64	11		35			36	33			81						33					28
		64														81	66	75					
	76			72	66	70	68	36		63	27	70	68	68					76				
75	64		66	72	64	75		66	75	83						37					69		28
80		72						90	66	75	83	63	64	64					66	66	71	76	71
									63	63													71
		70			72	70		63		63	66		35	35	35	37	35	84	69	69			64
				66		64						63			64	66					64		
				64				72	75	75	66			35		37							
				64	70			72	66		63							33			30		
				66	33			72	83	83	75	63	64		35					69	76	85	83
26	18			33		35					63										76		66

Seasonality: Systems, Strategies, and Signals by Jake Bernstein (New York: John Wiley & Sons, 1998).

*Trading Spreads and Seasonals* by Joe Ross (Ross Trading, 1995).

Seasonal Concepts in Futures Trading by Jacob Bernstein (New York: John Wiley & Sons, 1986).

In addition, many books include sections on seasonality within more general coverage of price patterns and strategies.

SEASONALITY AND THE STOCK MARKET

At the beginning of this chapter, Southwest Airlines was used as an example of a likely candidate for seasonality, and Figure 10.28 showed the seasonal pattern for AMR. Now we will look at the S&P for general patterns in the stock market. As with some of the other examples, S&P *e*mini futures will be used because that is the most popular contract, although SPYs could also be used to reflect prices closer to the cash index. Because futures will be used, the price differences will be the basis for the monthly patterns, rather than the returns. To minimize the effects of higher prices, the data from 1982 to 2010 will be separated into the first 19 years and the last 10 years. The results are shown in Figure 10.29.

We first see that the 19 years from 1982 to 2000 were far less volatile than recent years, with a clear decline during the summer months. In addition, the pattern from October through the end of November, or possibly December, has not changed. The theory is that investors come back from their summer holidays and move their money back into the market. They remove part of their investment at the end of the year for tax purposes. The old adage, "buy in May and go away" does not seem to be a good idea any more.
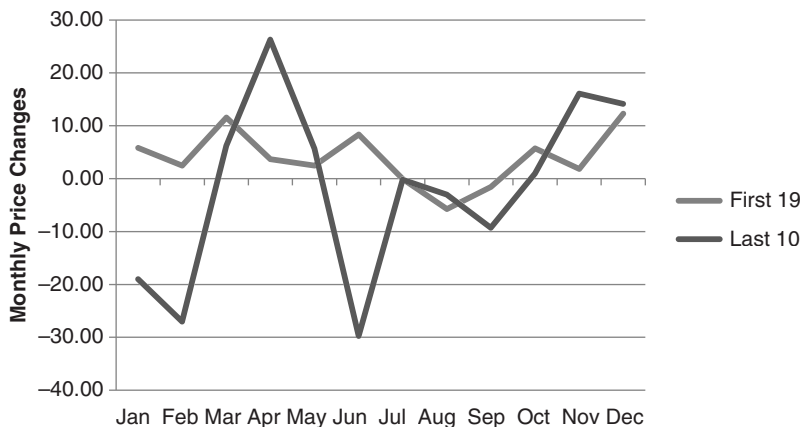
FIGURE 10.29   Seasonality of the S&P using back-adjusted *e*mini futures.
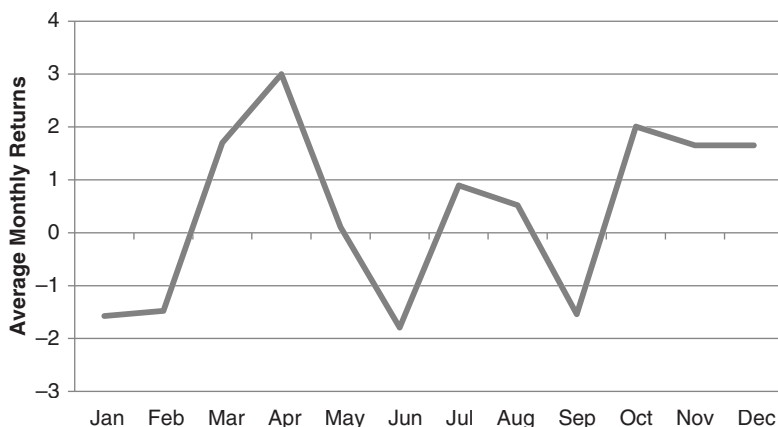
FIGURE 10.30 The DJIA average month-end cash prices, May 1998 through March 2011.

To show that it is not the discounting of futures prices that have caused this pattern, Figure 10.30 uses the cash Dow Industrials for the recent 12 years and averages the month-end returns. The pattern is very similar, peaking at the end of April, with a rally in October. The Dow is even less consistent than the S&P, but then it only has 30 stocks.

Most of all, these charts show that patterns have changed. Global factors, economic crises, arbitrage, and different participants all play a role. Only the most significant patterns will survive over time.

The Holiday Effect for Stocks

Arthur Merrill is acknowledged as the pioneer in seasonal timing of stocks, publishing his comprehensive work on this topic in 1966.[9] In his studies of price movement before and after major holidays, Merrill demonstrates a strong bullish tendency in advance of a holiday with a weak day immediately following. Remembering the bullish bias of the stock market (about 54% of all days were higher from 1897 to January 1964), Merrill's results are shown in Table 10.14. For trading, this would indicate the possibility of a sharp trending move prior to or throughout a holiday season.[10] Kaeppel is more specific, recommending buying on the close of the third day before an exchange holiday and selling on the close two days later (one day before the holiday).[11]

---

[9] Arthur A. Merrill, *Behavior of Prices on Wall Street* (Chappaqua, NY: Analysis Press, 1966). It is currently available from Analysis Press, 3300 Darby Road, #3325, Haverford, PA.

[10] An excellent summary of seasonal studies in the stock market can be found in Nelson F. Freeburg's reports, *Formula Research* (Memphis, TN: Formula Research, Inc.), 4745 Poplar Ave., Suite 307, 38117-4408. Material from this section was drawn from Vol. VI, nos. 10–11, December 2000 and August 2001. Also recommended is Steve Moore, "Playing the Seasonals," *Active Trader* 1, no. 8 (November 2000).

[11] Jay Kaeppel, "The Stock Market, the Calendar, and You," *Technical Analysis of Stocks & Commoditie*s (December 2002).

TABLE 10.14	Merrill's Holiday Results	
Period Tested	**Holiday or Holiday Period**	**% Upward Moves**
1897–1964	Day prior to all holidays	67.9%
1897–1964	Day after all holidays	50.8
1897–1964	Thanksgiving to New Year	74
1897–1964	July 4th to Labor Day	69
1931–1965	Before Christmas	74
1931–1965	Before New Year	75

Norman Fosback[12] confirmed Merrill's results by studying the returns based on a strategy that bought two days prior to a major holiday and exited on the day prior to the holiday, rather than waiting until the day following the holiday. Fosback's strategy yielded returns of 880% from 1928 through 1975 with 70% of the trades profitable, while holding long positions during the remaining days of the year would have lost 41%. Therefore, the net gains for the year were all associated with the few days prior to major holidays. In addition, the returns on holiday trading were accomplished by a risk exposure of only a few days each year. According to an updated study by Freeburg, the strategy used by Fosback is still profitable, but not as impressive. Freeburg also showed that the same strategy was profitable for U.S. 30-year Treasury bond futures as well as corporate bonds.

The Month-End Effect

It seems sensible that if you are willing to take the opposite position to the crowd, there are profits to be made. One such strategy takes advantage of month-end liquidation. Perhaps some investors close out positions before the end of the month in order to realize profits or losses; this is even more likely to happen at the end of a quarter or the calendar year. This effect could be helped by large funds that may exit positions to balance redemptions. Merrill, Fosback, Kaeppel, and Freeburg all confirm the success of a strategy that buys on the last day of the month, or the second-to-the-last day, then exits the trade on the fourth trading day of the new month. That takes advantage of large-scale, month-end liquidation, followed by resetting positions. Freeburg confirmed that this strategy was still viable for both the S&P Index and bonds, although not as profitable as when these markets had limited participation up to 1975. A similar study, where patterns are shown for *days of the month*, including more recent years, can be found in Chapter 15 for a selection of futures markets.

The Hirsch Strategy

One of the most popular of all seasonal strategies was the work of Yale Hirsch. He simply bought on the first day of November and sold on the last day of April, holding the position for six months.[13] In addition to satisfying certain tax requirements, this avoided the

---

[12] Norman G. Fosback, *Stock Market Logic* (Chicago: Dearborn Financial, 1998).

[13] Nelson Freeburg, see footnote 10.

period most traders consider the summer doldrums. The Hirsch strategy would have avoided the spectacular October losses as well as the disaster of 9/11/2001 but benefited from the subsequent recovery. Hirsch had discovered that virtually all gains in the stock market took place during those six months.

Hirsch's original strategy reinvested dividend income during the six months when you were not in the market. That advantage has diminished, but leveraged investing can replace that loss. Using futures, exchange traded funds, or leveraged funds available through Rydex and ProFunds can make up the difference.

The January Effect

Another price pattern may be seen in the action of the stock market during the month of January. There are many investors who are not as anxious to trade in and out of the market as professional managers and speculators. It is perfectly sensible to look for a pattern in the way many of the long-term investors set positions at the beginning of the year, the result of a reallocation of their portfolios, or resetting positions liquidated before the end of the year for tax reasons.

If January is a leading indicator of stock market movement throughout the rest of the year, a combination of patterns should be considered based on the few days immediately after the year begins, and the net market direction for the month of January.[14] Using the Dow Industrials as a stock market indicator, the January pattern was viewed from 1900 to 1989 in two parts, 1900–1937 and 1938–1989. The results are shown in Figure 10.31 for all years, 1900-1989, and Table 10.15 for the 52 years ending in 1989.
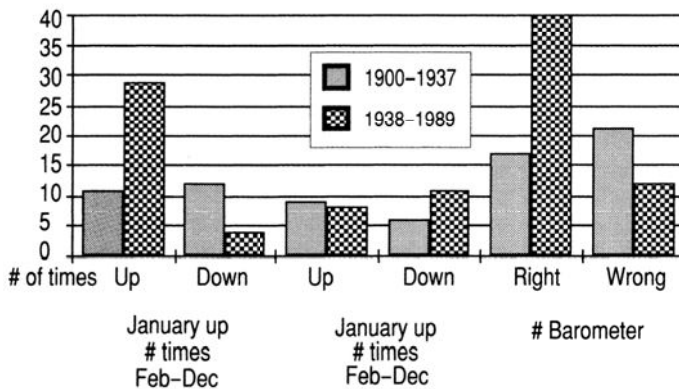
FIGURE10.31 Results of January patterns.

Source: Jay Kaeppel, "The January Barometer," *Technical Analysis of Stocks & Commodities* (July 1990).

---

[14]Jay Kaeppel, "The January Barometer: Myth and Reality," *Technical Analysis of Stocks & Commodities* (July 1990).

TABLE 10.15 The January Barometer Patterns, 1938–1989

	S&P 500		
Pattern	During First 5 Days of the Year	By the End of January	Expectations for Feb–Dec
1	Declines	Further decline	Bearish
2	Declines	Less decline, but loss for the month	Bearish
3	Declines	Gain for the month	Bullish
4	Advances	Further gains	Bullish
5	Advances	Less gain, but gain for the month	Bullish
6	Advances	Loss for the month	Bearish

Figure 10.31 shows that the early part of the century had no significant pattern. There were nearly the same number of bullish and bearish indications resulting in a larger number of incorrect predictions. The past 52 years are very different, with a much larger number of correct moves. One should note, however, that the ratio of up to down moves is more than 7:1. During the past 10 years it is likely that this pattern would continue to be successful if the prognosis was a bull market. To offset the market bias, the patterns shown in Table 10.14 can be used. The direction of the first five days, confirmed by a continuation in that pattern for the balance of January, is then followed by the same pattern for the year. In the case where the market changes direction after the first five days and nets a loss for the month, the pattern takes on the new direction for the balance of the year.

To bring the January pattern current to 2010, we used the monthly returns for the Dow previously used to create Figure 10.30. For the 12 years ending in 2010 there were 7 profitable years; that is, when January went up, then the rest of the year was up, and when January was down the rest of the year was lower. However, results were volatile, and the net was only 14 points. We can conclude that there is still a confirmation of the method, but it is very weak. Geopolitical events can overwhelm these patterns.

McGinley's January Indicator

John McGinley can confirm that January is significant for setting the bullish tone of the following year, but not for a bear market.[15] In his recent study, if each of the first five trading days of January are up, *buy* the S&P and hold for the entire year. If the end of January is up, then buy on February 1 and hold as well, although the end of January trade is not as good as the first five days.

McGinley states that, if the first 5 days are up at least 4%, the year has always been up. If we look at the most recent year in which this happened, 2010, the S&P Index (SPX)

---

[15]John R. McGinley, by email, January 2010. He is the editor of Technical Trends, a service that evaluates indicator accuracy. Contact John atjmcgoo@post.harvard.edu.

TABLE 10.16 McGinley's January Rule

	Dow Jones Industrial Average	S&P 500
Total Signals (since 1942)	44	40
No signal years	24	28
Correct signals	35	32
Number wrong	8*	7†
Percentage right	81.4!	82.1!
Chi Squared (highly significant #)	15.7	14.8
Years average was up	47	43
Years average was down	20	24
Percentage up	70	67

*1966, 1973, 1974, 1981, 1984, 1990, 2000, 2002
†1946, 1948, 1966, 1973, 1990, 1994, 2002
#>10.93 = greater than 999/1,000 it is not random

would have been entered at 1144.98 and ended the year at 1257.80, a gain of 9.85%. The next year, 2011, did not qualify. His performance summary is shown in Table 10.16.

A Chi-Squared reading this high (>10.93) indicates the odds are less than 0.01% that this pattern could occur by chance. Half of the 24 years with no Dow signal were down.

Risks of September and October

It is easy to point out the spectacular events of September and October: The tragedy of September 11, 2001; Black Monday on October 28, 1929; and the market crash of October 19, 1987. However, these months have other negative attributes that are not as noticeable. Freeburg has found 17 *ricochet rallies* in October since 1950. He defines a ricochet rally as a price rise of at least 5% within 10 days as measured from the S&P monthly low. June ran a distant second place with only 10 rallies. Therefore, while October may hold the record for the most volatile price moves and the most risk, it also has great opportunity for timing an entry into a new position.

While not as spectacular as October, September seems to capture consistency in underperformance in both recent years and throughout the past 100 years. Is this a remarkable coincidence, or is it likely to continue? The argument seems strong that investor behavior would continue; however, September can be a profitable month simply by chance without much effect on the probabilities.

COMMON SENSE AND SEASONALITY

Most comprehensive works on seasonality include not only agricultural markets but currencies, interest rates, and indexes. Very little is available on stocks. Common sense tells us that the major agricultural products follow a clear seasonal pattern based on Nature. This calendar cycle also applies to coffee, cocoa, and oranges, although competition

from South America has limited the extreme moves that resulted from a freeze in Florida's orange groves. But grain seasonality has been slowly changing as Brazil continues to improve its production of soybeans, Asian countries become a greater factor, and farmers build more on-site storage to take advantage of higher prices and the cost of carry. Cattle and hog prices, dependent upon the cost of feed, and still marketed more actively in the fall because of weather, show a strong seasonal trend even though they have a growth and feeding cycle that is not confined to a calendar year. Livestock patterns have adjusted to long-term competition from Australia and New Zealand.

These agricultural markets are reasonable candidates for seasonal analysis, but may require careful study to identify shifting patterns in competition and storage. What about Treasury bonds or the Japanese yen? Should we look for seasonal patterns in the financial markets? We can argue that certain listed companies and even market sectors are highly dependent upon seasonal business. However, can we say that the demand for money is higher in the summer than in the winter? Or that enough tourists convert their currencies to yen when they visit Japan during the summer that they effectively drive the value of the yen consistently higher? Are automobile exports from Japan stronger or weaker during a certain season, so that there is a predictable pattern? These scenarios are highly unlikely, and we have no assurance that cars will be sold off with the coming of winter in the way that grain must be harvested. In financial markets, most players can choose their time to act; they can hedge their commitments or wait for a better opportunity.

There are some stocks that clearly qualify for seasonality, such as airlines that service nonbusiness customers, resort hotels, cruise lines, perhaps even luggage manufacturers. Holiday goers tend to book in the spring for summer vacations. And it is possible that mortgages spike in the spring and are at lows in the winter simply because shoppers are not looking for houses when the weather is bad. The study of seasonality should be limited to those markets that depend on weather and seasons, on Mother Nature or consumer behavior, either directly or indirectly. If you find a pattern in a market that cannot be clearly explained by fundamentals, it is best to avoid it.

Being Too Specific about Targets

Nature is not precise, so selecting an *optimal day* of the year to enter a long position is not likely to work. Small shifts in fundamentals, such as the building of additional storage, allow farmers to change their selling habits slightly. The right day to buy or sell in 1975 is not likely to be the right day this year. Even weekly data may be too specific. Seasonal patterns are best seen using monthly data. The most we might expect is to know that the seasonal high in corn usually occurs in July but sometimes in June or August; the harvest lows are likely in September but could be in October or November. We must first be aware of the big picture and then study how the specific pattern develops each year. There are other timing tools, such as overbought and oversold indicators, that can help narrow the moment of opportunity within an expected window of time. If not, averaging into a position is always safe.

Cycle Analysis

The cycle is another basic element of price movement, along with the trend and seasonality, but as a mathematical problem it can be more difficult to evaluate and is often avoided. But there are many different types of cycles, from agricultural to presidential election, and many of them are simple to evaluate and can improve trading.

Cycles come in many forms—seasonality, production startup and shutdown, inventory or stocks, behavioral, and even astronomical. Seasonality is a special case of a calendar or annual cycle. Seasonality was covered in the previous chapter, and its special features are not considered here. Some of the cycles are clearly *periodic*, having regular intervals between peaks and valleys; others are more uniform in their *amplitude* or height but irregular in period. The most definitive and regular cycle remains the seasonal, which is determined by periodic physical phenomena, the changing of the year.

This chapter will discuss the major commodity and financial cycles that most likely result from business decisions, government programs, and long-term market characteristics and phenomena. Short-term cycles are usually attributed to behavior and will be covered in Chapter 15, Pattern Recognition.

There are a few important ways to find the cycle, the most common being *trigonometric curve fitting* and *Fourier (spectral) analysis*. Both will require a computer and will be explained in the following sections. John Ehlers introduced *Maximum Entropy Spectral Analysis* (MESA), which finds price cycles based on small amounts of data, at the same time avoiding some of the problems inherent in other methods. Examples of solutions will be included in the explanation of the methods and applications that follows. Computer programs that solve the trigonometric problems can be found on the Companion Website as well as in Appendix 3, along with additional examples. Commercial software examples will also be used to show the classic solutions.

CYCLE BASICS

The cycle, along with the trend and seasonality, comprise the three orderly components of price movement. The fourth is noise, which includes everything not accounted for in

the first three. To find any one component, we must remove the others. In earlier chapters we found that we can eliminate the trend by taking the first differences of the data; that is, subtracting the previous value from the current value. In the previous chapter on seasonality, we used the simple technique of subtracting a 1-year moving average (a 12-period average applied to monthly data) from the original price series to remove the seasonal pattern. Alternatively, statistical software will subtract this month's average price from that of 12 months ago, or today's daily price from that of 252 days ago in order to detrend the data. By finding the first differences and then subtracting the 1-year average, or by removing the 1-year differences, we are left with the cycle and the unaccountable price movement, which we call *noise*.

Even when the seasonal pattern is eliminated, most cycles are still based on the periodic effects in our Universe. After the 1-year orbit of our planet around the Sun, there is the 28-day lunar cycle; converted to business days, this gives the very familiar 20-day reference that remains overwhelmingly popular among all analysts (also corresponding to four weeks). Other planetary effects, which should by no means be discarded off-hand, can be found in Chapter 14 under the topic "Financial Astrology." The possibility cannot be eliminated that planetary motion may account for, besides seasonality, the effects of mass behavior, which can produce a consistent cycle that repeats with a fixed period.[1]

Cycles can be complex and difficult to see because there is often a combination of larger and smaller patterns, and cycles within cycles, all acting at the same time. Nevertheless, they exist, and they are real. The cycles that appear to be most important are either long-term or the sum of a number of subcycles that come together at peaks or valleys. This gives us a way to identify one point on a cycle; we must remember that, when the individual components are found, there may be a number of smaller patterns that cause this effect. Thinking about it as *harmonics*, just as in music, means that a smaller cycle is a fraction of the larger cycle, for example, its cycle length is ½, ⅓, ¼, . . . of the larger. When two cycles are *synchronized*, their peaks or valleys occur at the same time. Any price series can be decomposed into individual cycles, and represented as the sum of multiple cycles.

Observing the Cycle

Before selecting a market for cycle analysis, it is necessary to observe that a dominant cycle exists; it is also useful to know why it exists in order to avoid uncovering spurious patterns. This is most easily done for markets in which you can clearly identify the fundamental or industrial reasons for cycles. The basis for a cycle could be a pattern of holding inventory, the fixed time needed for breeding and feeding of livestock, seasonality, the time necessary for closing a mining operation then starting it up again, expansion or contraction of business based on disposable income, the effects of government interest rate policy, or other economic factors.

---

[1]See the section "The Moon," in Chapter 14.

The Cattle Cycle

Using cattle as an example, Figure 11.1a shows a clear 9- to 11-month cycle in futures prices[2] over a 6-year period from 1980 through 1985. The peaks and valleys vary by up to one month, making the pattern reliable for use as part of a long-term trading strategy. Although feedlots in the Southwest have made the supply of cattle more evenly distributed throughout the year, there are still a large number of ranchers in the North who send their cattle to market in the early fall to avoid the difficulties of feeding during a harsh winter. This causes generally lower prices in the Fall and higher prices in the mid-Winter when supplies are low.

A similar pattern can be seen more recently in Figure 11.1b. During the past six years the peaks of the cycle are consistently 12 months apart, although the valleys are not as consistent, most often coming within a few months after the peaks. The overall picture shows that cattle prices continue to have a clear cycle, driven by the fundamentals of production.

The Swiss Franc Cycle

The Swiss franc cycle (denominated as Swiss francs/U.S. dollars on Chicago's International Monetary Market) shown in Figure 11.2a is quite different.[3] There are two likely cycles: the primary one (shown using letters at the peaks and valleys) and a subcycle (marked with numbers). The subcycle ranges from 24 to 35 weeks with a 40% variance compared to 20% for cattle. Most important, the cycle in the Swiss franc cannot be attributed to any specific fundamental cause. There is certainly a long-term cycle based on the strength and weakness of the U.S. economy with respect to the Swiss economy, or the relative attractiveness of U.S. interest rates. There is also the general ebb and flow of the U.S. trade balance and, of course, investor behavior. Unlike cattle, these patterns do not need to be rigid.

Looking at Swiss franc prices from 1997 through 2002 there are obvious peaks and valleys that continue a cyclic pattern (see Figure 11.2b). Although they are crisp in appearance, the cycle now has an average period of about 38 weeks with a range from 30 to 52 weeks. The new cycle falls about midway between the periods of the previous primary and subcycles. Although the cycles seem clear, the change in period and the variance between cycle tops will make a systematic strategy difficult.

Basic Cycle Identification

A simple way to begin the search for major cycles is to look at a long-term chart, displayed as weekly rather than daily prices. The dominant half-cycle can be found by locating the obvious price peaks and valleys, then averaging the distance between them. A convenient tool for estimating the cycle length is the Ehrlich Cycle Finder.[4] Developed in

---

[2]Jacob Bernstein, "Cycle and Seasonal Price Tendencies in Meat and Livestock Markets," in Todd Lofton, ed., *Trading Tactics* (Chicago: Chicago Mercantile Exchange, 1986).

[3]Jacob Bernstein, *The Handbook of Commodity Cycles* (New York: John Wiley & Sons, 1982).

[4]More information and a cycle-finding tool can be found on www.stanehrlich.com.

(a)

(b)
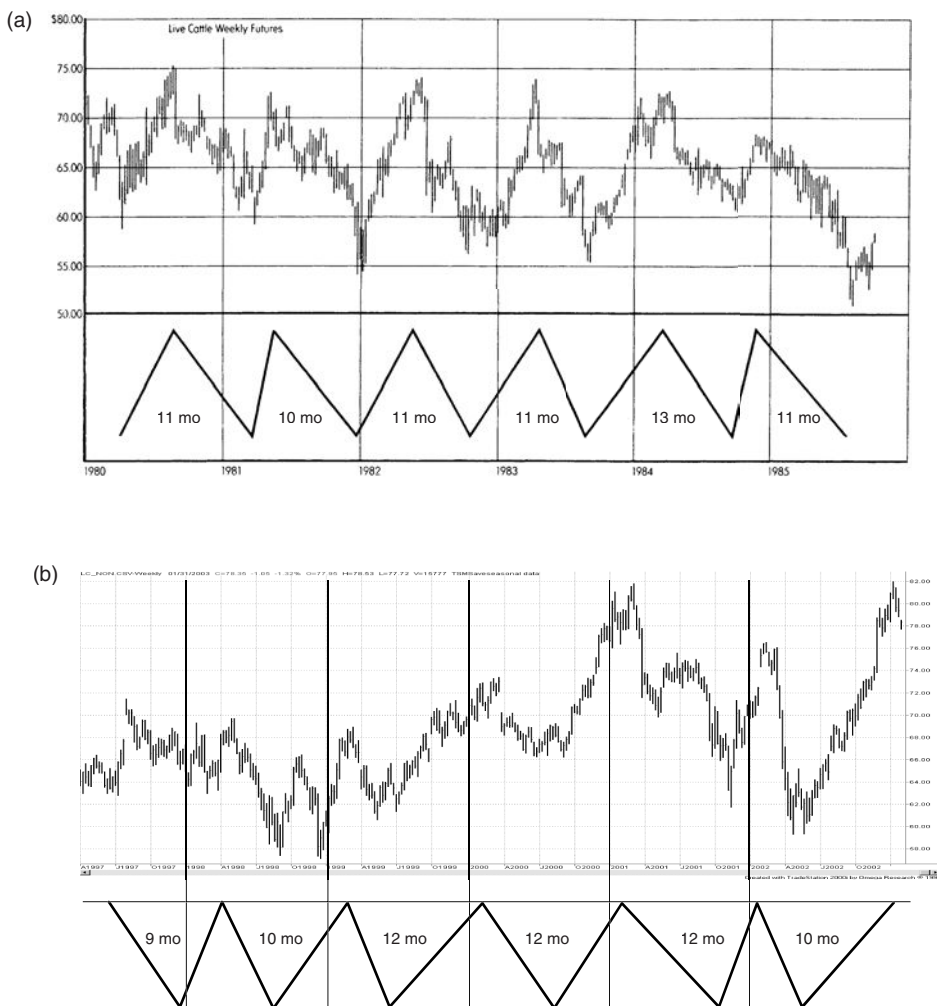
FIGURE 11.1 (a) 9- to 11-Month cycle in live cattle, 1980–1985 futures prices.
(b) The cattle cycle, 1997–2002.

1978, it is an expanding device with evenly spaced points, allowing you to align the peaks and valleys and to observe the consistency in the cycle. For finding a single pattern, it is just as good as some of the mathematical methods that follow. It is best to have at least eight cycle repetitions before concluding that you have a valid cycle.

Cycles can be obscured by other price patterns or market noise. Strong trends, such as the ones in Swiss francs (Figure 11.2a) or the seasonal movement of crops, may overwhelm a less pronounced cycle. Classic cycle identification requires that these noncycle factors first be removed by detrending and then by deseasonalizing. The resulting data will then be analyzed and the trend and seasonal factors added
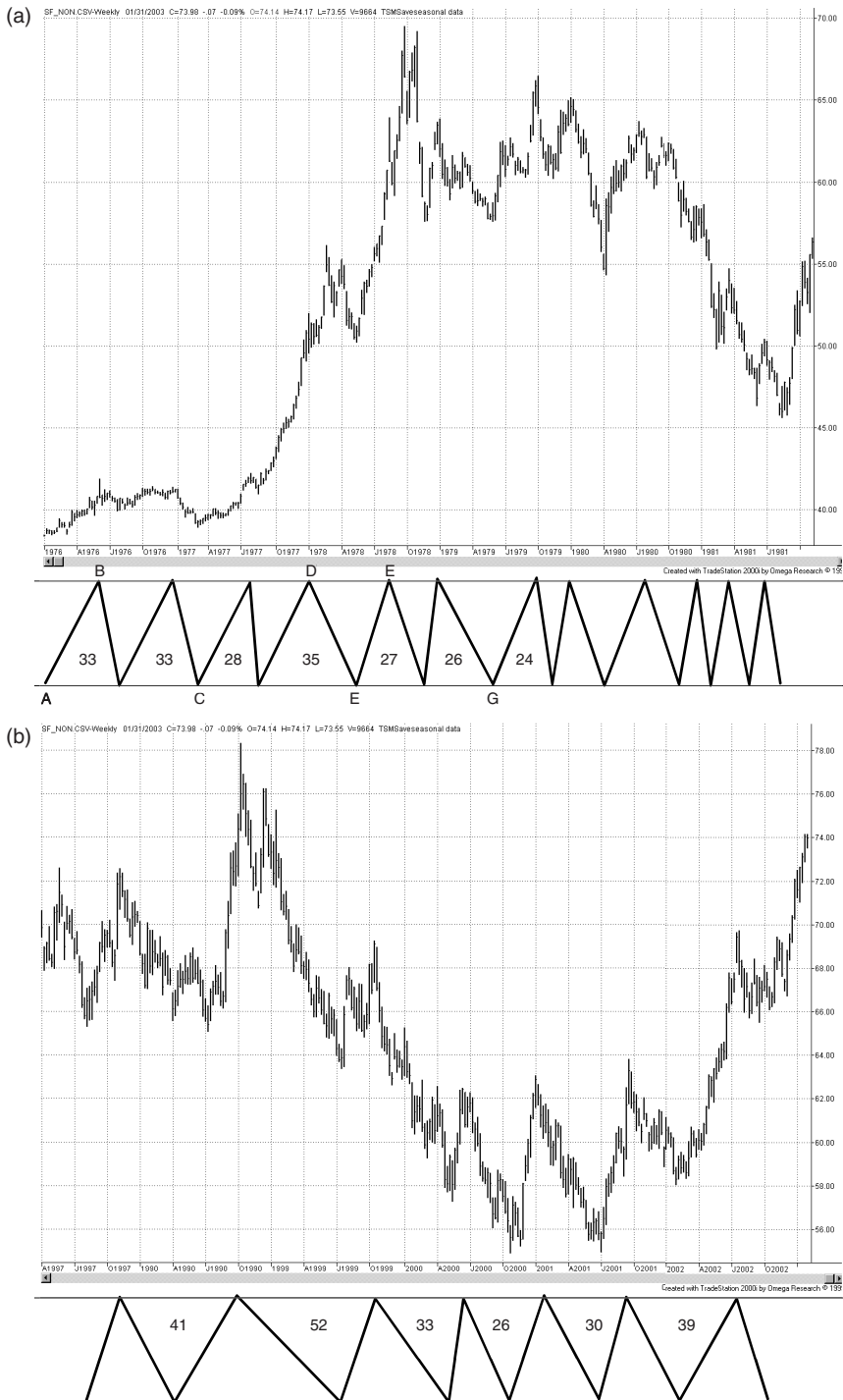
FIGURE 11.2 (a) Cycle in Swiss franc futures, 1975–1979. The lettered peaks and valleys show the choice for a primary cycle; the numbered peaks and valleys show a likely subcycle. (b) Cycle in Swiss franc futures, 1997–2002.

back once the cycle has been found. To find a subcycle, the primary cycle should be removed and a second cycle analysis performed on the data. This can be a tedious process. In order to bypass these steps, the methods that follow (trigonometric regression and spectral analysis) can locate the dominant cycle and subcycles at one time using an integrated process.

The Business Cycle

The global business cycle, as distinguished from industry cycles, is the result of macroeconomic events, such as recessions, inflation, and government economic policy. Figure 11.3, a product of the Princeton Economic Institute, shows that this cycle is about 8.6 years, or about 4 years from top to bottom in each cycle. Although this chart dates from 1997, it seems remarkably accurate in capturing the tech bubble that ended in 2000, the downturn that followed, ending in 2003, the rally preceding the 2008 subprime crisis, and the extreme fall of the market and the economy afterward. It shows the bottom of this cycle in 2011, which we all hope is true.

The Kondratieff Wave

Much of the popularity of cycles is due to the publicity of Nicolai Kondratieff's 54-year cycle, known as the *K-wave*, or more recently, the *long wave*. During its documented span from about 1780 to the present, it appears to be very regular, moving from highs to lows and back again. In Figure 11.4 the Kondratieff wave is shown with major events
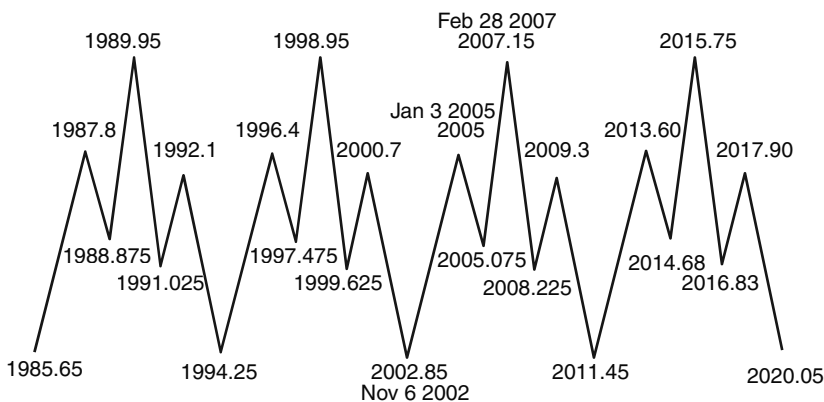
FIGURE 11.3   The 8.6-year business cycle.

Source: The Princeton Economic Institute, available on www.financialsense.com.

The Kondratieff wave

FIGURE 11.4   The Kondradieff Wave.

Source: Walker, Jeff, "What K-wave?" *Technical Analysis of Stocks & Commodities* (July 1990).

(particularly wars) that have contributed to its pattern.[5] With only three full cycles completed, it is difficult to tell if the overall trend is moving upwards, or whether the entire pattern is just a coincidence.

The forecast of the *K*-wave, shown in Figure 11.4, corresponds to a sharp decline in wholesale prices due at about the year 1990, the millennium's equivalent to the depression of the 1930s. In fact, the 1990s posted remarkable gains in the stock market, peaking at the beginning of 2000. According to the chart pattern, this peak should be followed by 10 to 20 years of downturn in the economy, in which case we are in the middle, having experienced a major correction in 2008. It should be noted that the peaks of the four waves are of different duration, 1870 being the shortest and the recent one in 2000 the longest.

Although we all accept the existence of an economic cycle, pinpointing the peaks and valleys is impractical. Even if the 54-year period varied only by 10%, we could be entering an investment position 5 years too soon or too late. Determining long-term cycles for any market has the same problem—the actual price pattern will never correspond exactly to the predicted peaks and valleys that most often come at regular intervals. Fortunately, there are other choices. Shorter-term cycles do not need to have the same constant period, and the way in which cycles interact with other strategy components will make them more flexible. However, some investors will want to keep this big picture

---

[5]Jeff Walker, "What K-Wave?" *Technical Analysis of Stocks & Commodities* (July 1990).

concept, both the business cycle and the Kondratieff wave, as a general guide to investment timing.

Presidential Election Cycle

Of all the events that move the market, the presidential elections have been the most consistent. The patterns stem from the motivation of the incumbent party to provide good economic news to the voters prior to the election year, and as far into the election year as possible. Stock market action during the election year is always more erratic, as parties battle over the value of each other's actions.

Typically, the year preceding the election (year 3 in the president's term) posts the strongest gains for the market, followed by a reasonably strong election year. (See Table 11.1.) Some analysts have been more specific by starting on October 1 of the previous year. The two years after the election show returns below average as the reality of politics reasserts itself and the new administration tries to implement campaign promises that turn out to be unpopular. More recently, it is only in the first year that the president can push for serious reform. Beginning in the second year, the mid-term elections of members of Congress become more important.

There is the additional possibility that there is an eight-year cycle that should be watched;[6] however, the eight-year period should be most informative if it represented only those years in which the same president was in office. Actions by a president who cannot be reelected are likely to be different from one who seeks another term; therefore, we should expect a different pattern. This can be made more intricate by studying the patterns preceding and following a change of party, all of which have a fundamental basis in the behavior of the political parties and the voters.

TABLE 11.1 The Presidential Election Cycle, 1912–1992, Based on the Percentage Returns of the Dow Jones Industrial Averages

Pre-election year	11.0%
Election year	7.0
Post-election year	4.7
Mid-term year	2.3
Average year	6.3%

Source: Adam White.

---

[6]Articles by Adam White, "The Eight-Year Presidential Election Pattern," *Technical Analysis of Stocks & Commodities* (November 1994); Arthur Merrill, "The Presidential Election Cycle," *Technical Analysis of Stocks & Commodities* (March 1992); and Michael J. Carr, "Get out the Vote and into Stocks," *Futures* (February 1996), all show very similar results for the four-year election pattern.

TABLE 11.2 Election Year Analysis for Years in Which the Stock Market Began the Year within 8% of the Previous 2-Year Highs

Year	1. Previous Year	2. First 2 weeks (1–10)	3. Primaries (10–83)	4. Pre-convention (83–161)	5. Pre-election (161–195)	6. Election to year-end (195–253)	7. (2) + (4) + (6)
1936	41.82	2.76	4.64	11.91	−0.62	5.85	20.52
1944	19.45	1.63	−0.84	9.27	−0.86	3.14	14.04
1952	16.15	1.60	−2.82	8.02	−2.49	5.47	15.10
1956	27.25	−1.78	5.42	3.16	−6.65	2.38	3.76
1960	8.48	−2.49	−5.75	2.94	−7.13	8.49	8.95
1964	18.89	1.79	4.44	3.03	2.52	0.06	4.88
1968	20.03	0.26	−0.10	1.44	4.39	4.25	5.95
1972	10.82	1.41	3.39	5.30	−1.78	4.88	11.59
1980	12.31	2.26	−5.42	18.57	−0.20	7.66	28.49
1984	17.53	1.27	−5.01	4.27	0.41	0.63	6.17
1992	26.30	0.80	−2.72	2.14	−0.23	5.87	8.82
Average	19.91	0.86	−0.44	6.37	−1.15	4.43	11.66

Source: Michael Carr, Logical Information Machines.

More sophisticated computer software, such as that provided by Logical Information Machines,[7] a Chicago firm, can produce a very interesting, closer view of how voters respond to election politics. Table 11.2 breaks the election year into seven periods between the key events for those years in which the stock market began the election year within 8% of its 2-year high price (*days* refer to business days):

1. The returns of the year preceding the election year.
2. The first 10 days of the new year, typically a strong period (days 0–10).
3. Through the State of the Union address and the primaries (days 10–83).
4. Waiting for the conventions (days 83–161).
5. Preelection blahs: the actual campaign (days 161–195).
6. The election to year-end reaction (days 195–253).
7. Combined periods (2) + (4) + (6).

Combining the three periods (2), (4), and (6), which have strong upward biases, gives consistently positive results. Even if the newly elected party fails to deliver on its campaign promises, traders could have already converted those marketing gimmicks into stock market profits.

---

[7]See www.lim.com.

TABLE 11.3	Updated Presidential Election Cycle Based on S&P Futures
Cycle	**S&P Total Returns**
Preelection	16.4
Election	69.5
1st year	(30.0)
2nd year	48.3

Presidential Cycle from 1983 to 2010

In our rapidly changing world, it is always interesting to see if the market reality continues to support expectations. In fact, using the S&P futures and calculating the year-end returns, the results (in Table 11.3) confirm our new expectations of the presidential cycle. There are moderately good returns in the preelection year, but excellent returns in the year of the election as all candidates and parties promise whatever is necessary to get elected.

Reality follows in the first year of office, when the president attempts to fulfill campaign promises but also takes this one opportunity for economic reforms that are likely to be unpopular, such as budget reductions and tax increases. A better year follows ahead of the mid-term elections, which have become a more important political event than in the past.[8]

UNCOVERING THE CYCLE

Before resorting to the highly mathematical methods for finding cycles, there are some simple approaches that may serve many traders. For example, if you believe that the dominant cycle has a 20-day period, then you simply create a new price series by subtracting the current data from a 20-day moving average. This removes the trend that may obscure the cycle. This is the same method used for removing seasonality, which subtracted the values of a one-year trend from the corresponding prices. Alternatively, you can take the 20-day differences ($p_t - p_{t-20}$), which effectively removes the 20-day trend.

Most oscillators, such as a stochastic or RSI, can also serve to identify a price cycle; however, if you want to see the peaks and valleys of a 20-day cycle, you will need to use a calculation period for the oscillators that is no more than 10 days.

Removing the Trend

The cycle can become more obvious by removing the price trend. While we traditionally only use one trendline to do this, the use of two trendlines seems to work very well in

---

[8]Gerald Appel states, "There is a clear election-year cycle, where the election year is +10%, year after +4.5%, 2 years before next –1.25%, and the year before +20%." *Technical Analysis: Power Tool for Active Traders* (Upper Saddle River, NJ: FT Prentice Hall, 2005), 94.

most cases.[9] First smooth the data using two exponential moving averages, where the longer average is half the period of the dominant cycle (using your best guess), and the shorter one is half the period of the longer one. Then create an MACD indicator by subtracting the value of one exponential trend from the other; the resulting *synthetic* series reduces the lag inherent in most methods while removing the trend.

Using Triangular Weighting

One method for enhancing the cycle is the use of *triangular weighting* instead of exponential smoothing. The weighting is *triangular* because it creates a set of weighting factors that are smallest at the ends and largest in the middle, and typically symmetric. You must first decide the calculation period and the weighting factor for the center price. For practical purposes, it is only necessary to give the center price the weight of 2.0. Because there needs to be a center price, the triangular weighting method will eliminate the oldest price if the calculation period is an even number.

The weighting factors begin with the value $2.0/(P/2)$, and increase by the same value. If the calculation period $P = 10$, the weighting begins at $t - P + 2$, or $t - 8$, eliminating the oldest value in order to have an odd number of prices. The weighting factors, w_i, are then 0.4, 0.8, 1.2, 1.6, 2.0, 1.6, 1.2, 0.8, and 0.4. The triangular average is

$$TMA_t = (w_1 \times P_{t-P+2} \times w_2 \times P_{t-P+3} + \ldots + w_{P-1} \times P_t) / P$$

Enhancing the cycle requires that you calculate two triangular averages, one of which is half the period of the other, then take the difference of the two. The smooth curve of the *triangular MACD* in Figure 11.5 shows the enhanced cyclic pattern of IBM
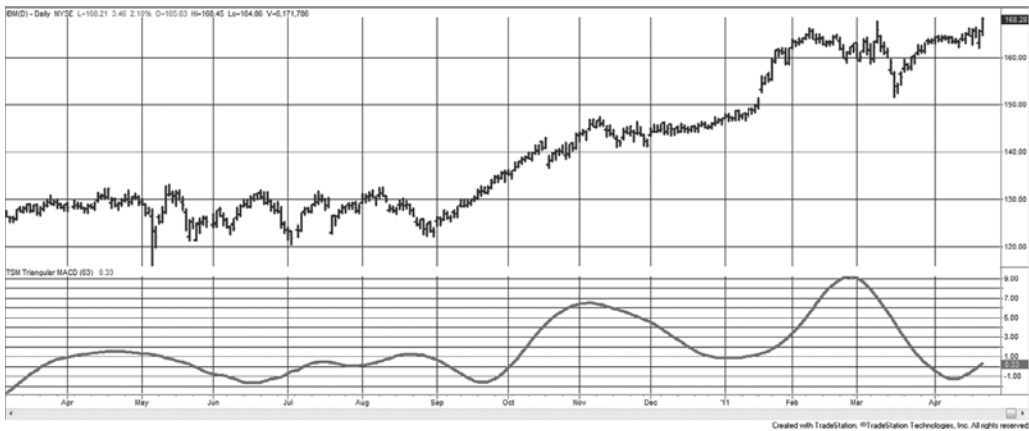
FIGURE 11.5 A 20-10 triangular MACD applied to IBM.

---

[9] In his article "Finding Cycles in Time Series Data," *Technical Analysis of Stocks & Commodities* (August 1990), A. Bruce Johnson credits John Ehlers for his work in the use of two exponential trends. See John Ehlers, "Moving Averages, Part 1" and "Moving Averages, Part 2," *Technical Analysis of Stocks & Commodities* (1988).
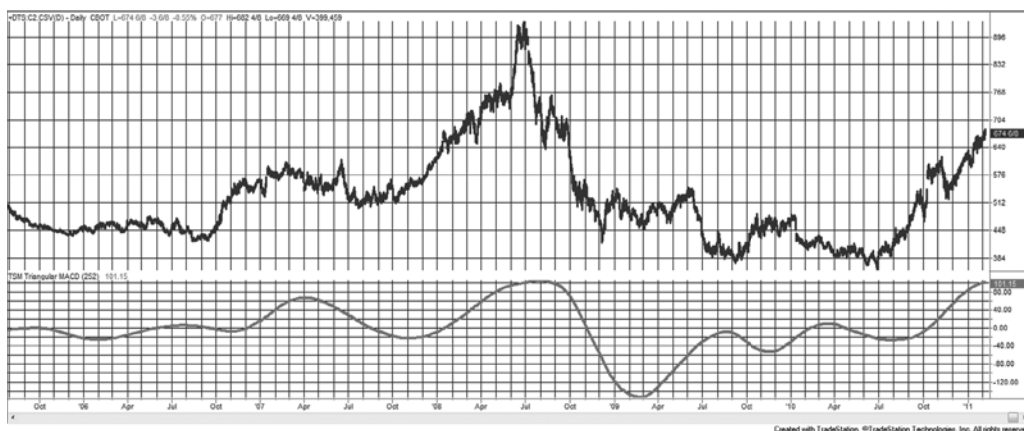
FIGURE 11.6   A 252-126 triangular MACD applied to back-adjusted corn futures.

(shown for one year) based on 63 days (about the number of days in a calendar quarter) under the assumption that cycles are likely to be related to periodic earnings releases. Figure 11.6 uses a full year, 252 days, as the primary cycle, with back-adjusted corn prices covering more than five years.

While this method creates a smooth representation of cyclic movement, it has the characteristics of a momentum indicator because the peaks and valleys are not quite evenly spaced and the amplitude of the cycles varies considerably. However, the smoothness of the indicator allows you to anticipate the major changes in the direction of prices. A program, *TSM Triangular MACD*, is available on the Companion Website.

Terminology

Before getting technical about the measurement and calculation of cycles, there are a few terms that describe most of the concepts discussed throughout this chapter. Note that the use of *wave* and *cycle* are interchangable.

> *Cycle or wave.* A recurring process that returns to its original state.
> *Amplitude* (*a*). The height of the wave from its horizontal midpoint (the *x*-axis).
> *Period* (*T*). The number of time units necessary to complete one wavelength (cycle).
> *Frequency* (*ω*). The number of wavelengths that repeat every 360°, calculated as $\omega = 1/T$.
> *Phase.* A measurement of the starting point or offset of the cycle relative to a benchmark wave.
> *Phase angle.* Locates the position within the cycle measured as the minute hand of a clock moving clockwise, where 0° is three o'clock.
> *Left and right translation.* The tendency for a cycle peak to fall to the left or right of the center of the cycle.

Trigonometric Price Analysis

Although there is statistical software that can find the cycle period with little trouble, some analysts still find it useful to know just how it can be done. Cycles can be found using the trigonometric functions *sine* and *cosine*. These functions result in what are called *periodic waves* because they repeat every 360° or 2π (2 pi) radians, where $\pi = 3.141592$. Because radians can be converted to degrees using the relationship

$$1 \text{ degree } = \frac{2\pi}{360}$$

all the work that follows will be in degrees.

A simple sine wave fluctuates back and forth from +1 to −1 (0, +1, 0, −1, 0) for each cycle (one wavelength) as the degrees increase from 0° to 360° (see Figure 11.7). To relate the wavelength to a specific distance in boxes (or days, on graph paper), simply divide 360° by the number of boxes in a full wavelength, resulting in the box size (in degrees). For example, a 100-box cycle would give a value of 3.6° to each box or day. The wavelength can be changed to something other than 360° by using the frequency, ω, as a multiplier of the angle of the sine wave, ϕ,

$$\sin \omega\phi$$

If $\omega > 1$, the frequency increases and the wavelength shortens to less than 360°; if $\omega < 1$, the frequency decreases and the wavelength increases. Because ω is the frequency, it gives the number of wavelengths in each 360° cycle. To change the *phase* of the wave (the starting point), the value b is added to the angle

$$\sin (\omega\phi + b)$$
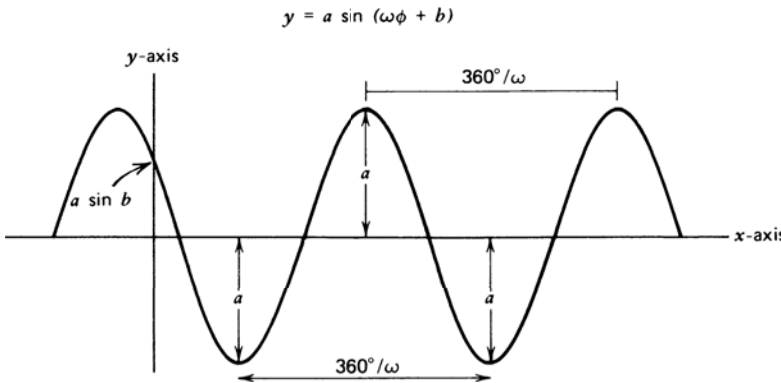
FIGURE 11.7 Sinusoidal (sine) wave.
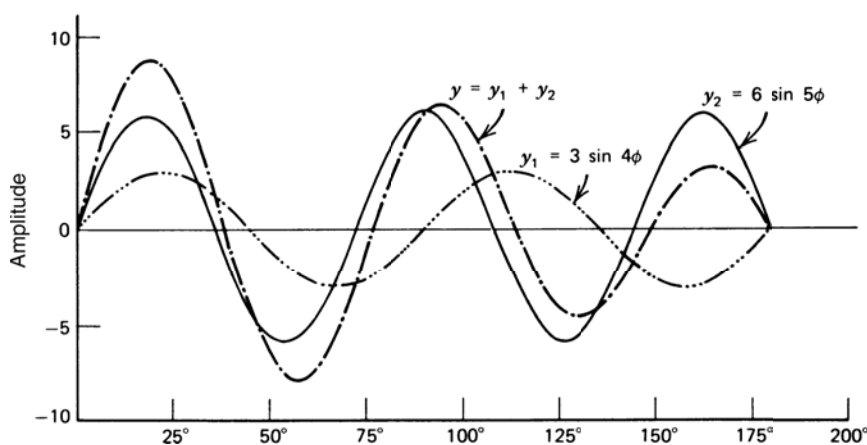
FIGURE 11.8   Compound sine wave.

If b is 180°, the sine wave will start in the second half of the cycle; the phase value b serves to shift the wave to the left. The *amplitude* can be changed by multiplying the resulting value by a constant a. Because the sine value ranges from +1 to −1, the new range will be +a to −a (Figure 11.8). This is written $a \sin(\omega\phi + b)$.

There are few examples of price movement that can be represented by a single wave; therefore, two sine waves (or more) must be added together to form a *compound wave:*

$$y = a_1 \sin(\omega_1\phi + b_1) + a_2 \sin(\omega_2\phi + b_2)$$

Each set of characteristic variables, a_1, ω_1, b_1, and a_2, ω_2, b_2, can be different, but both waves are measured at the same point ϕ at the same time. Consider an example that lets the phase constants b_1 and b_2 be zero. Then

$$y_1 = 3 \sin 4\phi$$
$$y_2 = 5 \sin 6\phi$$
$$y = y_1 + y_2$$

Figure 11.8 shows the individual regular waves y_1 and y_2, and the compound wave y over the interval 0° to 180°. Note that both y_1 and y_2 began the normal upward cycle at 0°; by 180°, however, they are perfectly out of phase. During the next 180°, the two waves come back into phase.

When combining periodic waves, it is useful to know the maximum and minimum amplitude of the resulting wave. Because the peaks of the two elementary waves do not necessarily fall at the same point, the maximum amplitude of either wave may not be reached. A mathematical technique called *differentiation* is used to find the maximum

and minimum amplitudes. The first derivative, with respect to angle ϕ, is written $dy/d\phi$ or y', where y is the formula to be differentiated. The rules are:

$$\frac{d}{d\phi}(\sin \phi) = \cos \phi; \quad \frac{d}{d\phi}(\cos \phi) = -\sin \phi$$

$$\frac{d}{d\phi}(\sin \omega\phi) = \omega \cos \omega\phi$$

$$\frac{d}{d\phi}(\sin (\omega\phi + b)) = \omega \cos (\omega\phi + b)$$

$$\frac{d}{d\phi}(a_1 \sin (\omega_1\phi + b_1) + a_2 \sin (\omega_2\phi + b_2))$$

$$= a_1\omega_1 \cos(\omega_1\phi + b_1) + a_2\omega_2 \cos(\omega_2\phi + b_2)$$

Applying this method to the previous example,

$$y = 3 \sin 4\phi + 6 \sin 5\phi$$

$$\frac{dy}{d\phi} = y' = 12 \cos 4\phi + 20 \cos 5\phi$$

The points of maximum and/or minimum value occur when $y' = 0$. For $y'_1 = 12 \cos 4\phi$, the maxima and minima occur when $4\phi = 90°$ and 270 ($\phi = 22\frac{1}{2}°$ and $67\frac{1}{2}°$) (Figure 11.8). For $y'_2 = 30 \cos 5\phi$, the maximum and minimum values occur at $5\phi = 90°$ and $270°$ ($\phi = 18°$ and $54°$). It must be pointed out that the first derivative identifies the location of the extreme highs and lows, but does not tell which one is the maximum and which is the minimum. The second derivative, y'', calculated by taking the derivative of y', is used for this purpose:

If $y'(x) = 0$ and $y''(x) > 0$, then $y(x)$ is a minimum.

If $y'(x) = 0$ and $y''(x) < 0$, then $y(x)$ is a maximum.

Then, $y_1 = 22\frac{1}{2}°$ and $y_2 = 18°$ are maxima and $y_1 = 67\frac{1}{2}°$ and $y_2 = 54°$ are minima.

Anyone interested in pursuing the analysis of extrema will find more complete discussions in a text on calculus. Rather than concentrating on these theoretical aspects of curves,[10] consider a practical example of finding a cycle in the price of scrap copper, shown in Table 11.4 and charted in Figure 11.9. The price peaks seem evenly spaced, occurring at mid-1966, January 1970, and January 1974, about four years apart. The solutions to these problems are tedious; therefore, calculations will be performed using the Fortran programs, *TSM Single Frequency Trigonometric Regression* and *TSM 2-Frequency Trigonometric Regression*, given on the Companion Website. These can be easily adapted to any programming language.

---

[10]A more detailed presentation of trigonometric curve fitting can be found in Claud Cleeton, *The Art of Independent Investing* (Upper Saddle River, NJ: Prentice-Hall, 1976), Chapter 8. The material covered in this section is carried further in that work.

TABLE 11.4 Dealer's Buying Price, No. 2 Heavy
Copper Scrap at New York*

Average Quarterly Price (cents per pound)

Year	1st	2nd	3rd	4th
1963	22.12	22.46	22.17	22.00
1964	23.18	24.56	25.57	30.59
1965	28.23	33.77	35.90	40.05
1966	46.22	51.48	40.76	40.16
1967	36.51	29.30	30.36	36.42
1968	39.75	30.07	29.08	32.13
1969	38.94	42.95	43.38	46.23
1970	47.70	46.98	35.78	27.35
1971	25.40	29.45	27.15	28.48
1972	32.74	33.53	30.01	29.25
1973	36.82	45.07	55.13	65.51
1974	66.56	70.06	47.30	35.62
1975	32.06	31.46	35.75	36.46
1976	38.22	43.24	45.46	38.96
1977	37.08	38.72	34.01	33.00
1978	35.07	40.23	41.63	44.95
1979	51.12	63.71	59.56	63.38

*Based on prices from the American Metal Market.

The results obtained by using actual copper prices will not be as clear as using ficti-
tious data. It is important to be able to understand the significance of practical results
and apply them effectively. Older data is used under the assumption that business and
economic cycles were clearer before globalization and geopolitical crises.
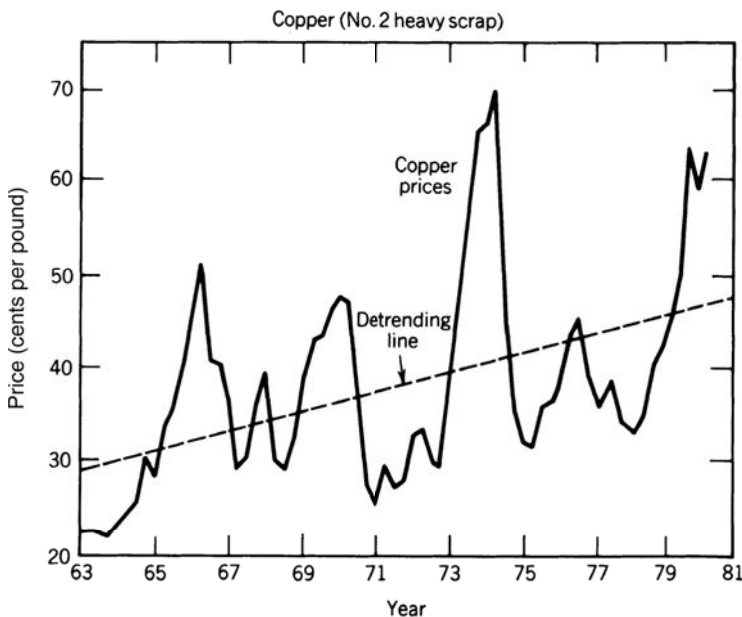
FIGURE 11.9 Copper prices 1963–1979.

Because trigonometric curves fluctuate above and below a horizontal line of value zero, the first step is to detrend the data using the least-squares method. This results in the equation for a straight line representing the upward bias of the data. The value of the regression line is then subtracted from the original data to produce copper prices that vary equally above and below the line from positive to negative values.

The straight line $y = a + bx$, which best represents the trend, can be found by solving the least-squares equations (also see Chapter 6):

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$a = \frac{1}{N}(\sum y - b \sum x)$$

To do this, let x be the date and y be the price on that date. For convenience, instead of letting $x = 1967, 1967\frac{1}{4}, 1967\frac{1}{2}, \ldots$, let $x = 1, 2, 3, \ldots$. The solution, using the program on the Companion Website (or the hand-calculation method in Chapter 6, is

$$y = 28.89 + 0.267x$$

Figure 11.9 shows the original copper prices and the regression line. The original prices can now be detrended using the equation above, subtracting the line values from the corresponding prices. Complete step-by-step results for this example can be found with the programs on the Companion Website. The detrended data is now used in the general trigonometric single-frequency wave:

$$y_t = a \cos \omega t + b \sin \omega t$$

The variable t replaces ϕ in order to consider the angle in integer units rather than in degrees. This will be more convenient to visualize and to chart.

To find the frequency ω, it is necessary to first solve the equation

$$\cos \omega - \tfrac{1}{2}\alpha = 0$$

using the system of equations,

$$\alpha y_2 = y_1 + y_3$$
$$\alpha y_3 = y_2 + y_4$$
$$\vdots$$
$$\alpha y_{n-1} = y_{n-2} + y_n$$

This is expressed as a summation (similar to least-squares) in which the values for c and d must be found:

$$\alpha \sum c^2 = \sum cd$$

where $\quad c = y_n$

$\qquad\quad d = y_{n-1} + y_{n+1}$

Summing the detrended values c^2 and cd gives $\Sigma c^2 = 6338.4$ and $\Sigma cd = 9282.2$, resulting in $\alpha = 1.464$. The value for α is substituted into the intermediate equation and solved for the frequency ω:

$$\cos \omega - \tfrac{1}{2}(1.464) = 0$$
$$\cos \omega = 0.732$$
$$\omega = 42.9$$

The period T is $360/42.9 = 8.4$ calendar quarters. The last step in solving the equation for a single frequency is to write the normal equations:

$$a\Sigma \cos^2 \omega t + b\Sigma\Sigma \cos \omega t \sin \omega t = \Sigma y_t \cos \omega t$$
$$a\Sigma \sin \omega t \cos \omega t + b\Sigma \sin^2 \omega t = \Sigma y_t \sin \omega t$$

and solve for a and b, where $t = 1, \ldots, 40$, and $\omega = 42.9$. As in the other solutions, a computer program is best for finding the sums (using detrended data) necessary to solve the equations. The sums are

$$a\Sigma \cos^2 \omega t \qquad \Sigma \sin \omega t \cos \omega t \qquad \Sigma y_t \cos \omega t$$
$$\Sigma \sin^2 \omega t \qquad \Sigma y_t \sin \omega t$$

Then, a and b can be found by substituting in the following equations:

$$b = \frac{\sum y_t \sin \omega t \sum \cos^2 \omega t - \sum y_t \cos \omega t}{\sum \sin^2 \omega t \sum \cos^2 \omega t - \sum \cos \omega t \sin \omega t}$$

$$a = \frac{\sum y_t \cos \omega t - b \sum \cos \omega t \sin \omega t}{\sum \cos^2 \omega t}$$

The results $a = -0.603$ and $b = 1.831$ give the single-frequency curve as:

$$y_t = -0.603 \cos 42.9t + 1.831 \sin 42.1t$$

Taking $t = 1$ to be 1967 and $t = 68$ to be 1979¾ and adding back the trend, the resulting periodic curve is shown in Figure 11.10.

The single-frequency curve shown in Figure 11.10 matches seven out of the eight peaks in copper; however, it is not much more than could have been done using the Ehrlich Cycle Finder. A single-frequency curve can be created simply by identifying the most dominant peaks, averaging the distance (period), and applying the single-frequency formula.

2-Frequency Trigonometric Regression

The combination of more than one set of sine and cosine waves of varying amplitudes and frequencies creates a better fit than a single-frequency solution. This is analogous to
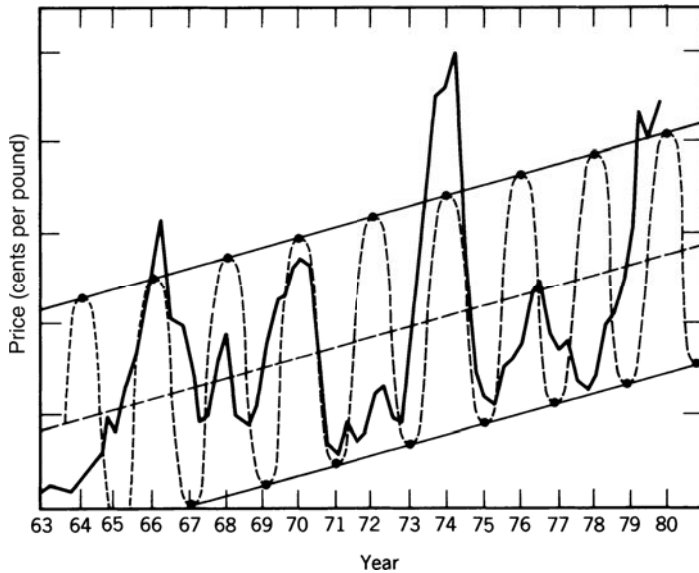
FIGURE 11.10 Copper prices 1963–1979 with single frequency copper cycle manually scaled to approximate amplitude.

the use of a second-order (curvilinear) solution instead of the first-order linear. The equation for the 2-frequency cycle is

$$y_t = a_1 \cos \omega_1 t + b_1 \sin \omega_1 t + a_2 \cos \omega_2 t + b_2 \sin \omega_2 t$$

To find the results of this complex wave, apply the same techniques used in the single-frequency approach to the detrended copper data. The algebra for solving this problem is an expanded form of the previous solution, and the use of a computer is a requirement (available on the Companion Website and described in Appendix 3). The frequencies ω_1 and ω_2 are found by solving the quadratic equation:

$$2x^2 - \alpha_1 x - (1 + \alpha_2 / 2) = 0$$

where $x = \cos \omega$, using the standard formula:

$$x = \frac{\alpha_1 \pm \sqrt{\alpha_1^2 + 8(1 + \alpha_2 / 2)}}{4}$$

The same least-squares method as before can be used, derived from the general form:

$$\alpha_1 (y_n + y_{n+2}) + \alpha_2 y_{n+1} = y_{n-1} + y_{n+3}$$

The least-square equations for finding α_1 and α_2 are:

$$\alpha_1 \Sigma c^2 + \alpha_2 \Sigma cd = \Sigma cp$$

$$\alpha_1 \Sigma cd + \alpha_2 \Sigma c^2 = \Sigma dp$$

where $c = y_n + y_{n+2}$
 $\quad\quad d = y_{n+1}$
 $\quad\quad p = y_{n-1} + y_{n+3}$

These equations can be solved for α_1 and α_2 using:

$$\alpha_2 = \frac{\sum dp \sum c^2 - \sum cp \sum cd}{\sum d^2 \sum c^2 - (\sum cd)^2}$$

$$\alpha_1 = \frac{\sum cp - \alpha_2 \sum cd}{\sum c^2}$$

Then, ω_1 and ω_2 are calculated from the two solutions x_1 and x_2 of the quadratic equation. The next step is to solve the normal equations to find the amplitudes a_1, b_1, a_2, and b_2:

$$a_1 \Sigma \cos^2 \omega_1 t + b_1 \Sigma \cos \omega_1 t \sin \omega_1 t + a_2 \Sigma \cos \omega_1 t \cos \omega_2 t$$
$$+ b_2 \Sigma \cos \omega_1 t \sin \omega_2 t = \Sigma y_t \cos \omega_1 t$$
$$a_1 \Sigma \sin \omega_1 t \cos \omega_1 t + b_1 \Sigma \sin^2 \omega_1 t + a_2 \Sigma \sin \omega_1 t \cos \omega_2 t$$
$$+ b_2 \Sigma \sin \omega_1 t \sin \omega_2 t = \Sigma y_t \sin \omega_1 t$$
$$a_1 \Sigma \cos \omega_2 t \cos \omega_1 t + b_1 \Sigma \cos \omega_2 t \sin \omega_1 t + a_2 \Sigma \cos^2 \omega_2 t$$
$$+ b_2 \Sigma \cos \omega_2 t \sin \omega_2 t = \Sigma y_t \cos \omega_2 t$$
$$a_1 \Sigma \sin \omega_2 t \cos \omega_1 t + b_1 \Sigma \sin \omega_2 t \sin \omega_1 t + a_2 \Sigma \sin \omega_2 t \cos \omega_2 t$$
$$+ b_2 \Sigma \sin^2 \omega_2 t = \Sigma y_t \sin \omega_2 t$$

Once the sums are obtained, the final step is to create a 4×5 matrix to solve the four normal equations for the coefficients a_1, b_1, a_2, and b_2. When plotting the answer, it will be best to plot the original 2-frequency equation in its component forms as well as in combination:

$$y_t' = a_1 \cos \omega_1 t + b_1 \sin \omega_1 t$$
$$y_t'' = a_2 \cos \omega_2 t + b_2 \sin \omega_2 t$$
$$y_t = y_t' + y_t''$$

where $a_1 = 3.635$
 $\quad\quad b_1 = -0.317$
 $\quad\quad a_2 = -0.930$
 $\quad\quad b_2 = 0.762$

The solution to the 2-frequency problem gives the following values:

$$\alpha_1 = 0.535, \quad\quad x_1 = 0.830$$

and

$$\alpha_2 = 0.133, \quad\quad x_2 = -0.764$$
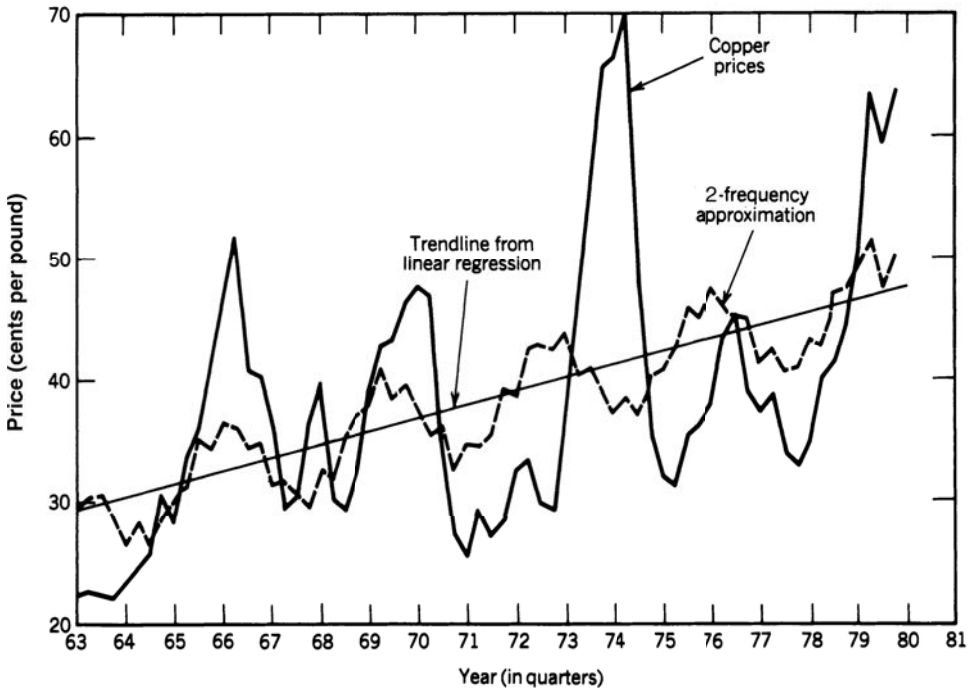
FIGURE 11.11 Two-frequency trigonometric approximation.

and finally the frequencies:

$$\omega_1 = 33.9 \qquad \text{and} \qquad \omega_2 = 139.8$$

correspond to 10.6 and 2.6 calendar quarters (Figure 11.11).

Fourier Analysis: Complex Trigonometric Regression

Developed by the French mathematician Jean Baptiste Joseph Fourier, *Fourier analysis* is a method of complex trigonometric regression, which expresses any data series as a series of sine and cosine waves of the same type as discussed in the previous section.

It should be understood that *Fourier analysis*, or the more popular and efficient *Fast Fourier Transform* (*FFT*), is not reliable unless the data is stationary, that is, it does not change in a disorderly way over time. Agricultural products are most likely to be candidates for this method, but stocks that have changed in price and changed in structure (such as those with new acquisitions that offer diversification), would not be candidates. In this section FFT will be applied to both corn, to see if the results are similar to the seasonality results from the previous chapter, and Southwest Airlines, which represents a company dependent on seasonal travel and without diversification. For best results,

it may be necessary to remove the rate of inflation or the USD exchange rate from the series to make it stable.

Basic Fourier Calculations

Assuming that there is a cycle and that there are N data points in each repetition of this cycle, the Fourier method of analysis shows that the N points lie on the regression curve:

$$y_i = 1 + \sum_{k=1}^{(N/2)} \left(u_k \cos \frac{2\pi ki}{(N/2)} + v_k \sin \frac{2\pi ki}{(N/2)} \right)$$

where the regression coefficients u_k and v_k are given by:

$$u_k = \frac{1}{(N/2)} \sum_{i=1}^{N} y_i \cos \frac{2\pi ki}{(N/2)}, k = 1, 2, \cdots, \frac{N}{2}$$

$$v_k = \frac{1}{(N/2)} \sum_{i=1}^{N} y_i \sin \frac{2\pi ki}{(N/2)}, k = 1, 2, \cdots, \frac{N}{2}$$

$$v_{N/2} = 0$$

It is important to see that the mean of all the points on one cycle is equal to the value 1. The N values of y_i will have the property

$$\sum_{i=1}^{N} y_i = N$$

Applying the Fourier series to the seasonal component will help clarify this method. Seasonal data form the most obvious cycle. Using average monthly prices, detrend the data to avoid letting the trend overwhelm the cycle, letting $N = 12$. It is also known that seasonally adjusted prices will vary about the mean; hence the weighting factors will have the same property as the above equation. With this information, the trigonometric curve that approximates the seasonality can be generated and compared with the results of other methods.[11]

Excel provides Fourier analysis as an analytic *add-in* under *Tools/Data Analysis*. Instructions and examples will follow in the section "Using Excel's Fourier Analysis."

Spectral Analysis

Derived from the word *spectrum*, spectral analysis is a statistical procedure that isolates and measures the cycles within a data series. The specific technique used is the Fourier series as previously discussed, although other series can be used.

---

[11]A continuation of this development can be found in Warren Gilchrist, *Statistical Forecasting* (London: John Wiley & Sons, 1976), 139–148; a more theoretical approach is to be found in C. Chatfield, *The Analysis of a Time Series: Theory and Practice* (London: Chapman and Hall, 1975), Chapter 7.

When studying the cycles that comprise a data series, it is important to refer to their *phase* with respect to each other. Phase is the relationship of the starting points of different cycles. For example, if one cycle has the same period as another but its peaks and valleys are exactly opposite, it is 180° *out of phase.* If the two cycles are identical in *phase*, they are *coincident.* Cycles with the same period may lead or lag the other by being out of phase to various degrees.

A tool used in spectral analysis to visualize the relative significance of a series' cyclic components is the *periodogram.* Weighting the cyclic components in the periodogram will yield the more popular *spectral density* diagram, which will be used to illustrate the results of the spectral analysis. *Density* refers to the frequency of occurrence. Figures 11.12a and 11.12b show the spectral density of a series composed of three simple waves (*D* is the Fourier series made up of waves *A*, *B*, and *C*).[12] The cycle length, shown at the bottom of the spectral density chart, corresponds exactly to the cycle length of the component waves *A*, *B*, and *C*. The spectral density, measured along the left side of Figure 11.12b, varies with the amplitude squared of the cycle and the magnitude of the noise, as well as random price movements, which tends obscures the cycle. In Figure 11.12b, the result is based on a series composed of only three pure waves. Had there been noise, of the same magnitude as the underlying cycle amplitude, those cycles identified by the spectral analysis would have been completely obscured. Readers who have studied ARIMA will recognize the similarity between the spectral density and the correlogram.

As in trigonometric regression analysis, the other basic price components can distort the results. A noticeable trend in the data must be removed, or it may be interpreted as the dominant cycle. The familiar methods of first differencing or linear regression can be used to accomplish this. The seasonal component is itself a cycle and does not need to be removed from the series. Because spectral analysis identifies both the seasonal and cyclic component, the success of the results will depend on the strength of these waves compared to the noise that remains. In applying this technique to real data, it would not be surprising to see the results demonstrated in Figure 11.13. Three subcycles of length 10, 20, and 40 days are shown as part of a 250-day (seasonal) cycle. Notice that, as the cycle lengthens, the width of the spectral density representation widens. This does not mean that the wider peaks are more important.

The trader is most interested in those cycles with greater spectral density, corresponding to a larger price move. The minimum amount of data necessary to find these cycles must include the full cycle that might be identified. For example, to see any seasonal pattern, a minimum of 12 months is needed. More data are better when using spectral analysis to confirm the consistency of the cycle. A single year is not enough to support any seasonal findings.

---

[12] William T. Taylor, "Fourier Spectral Analysis," *Technical Analysis of Stocks & Commodities* (July/August 1984).
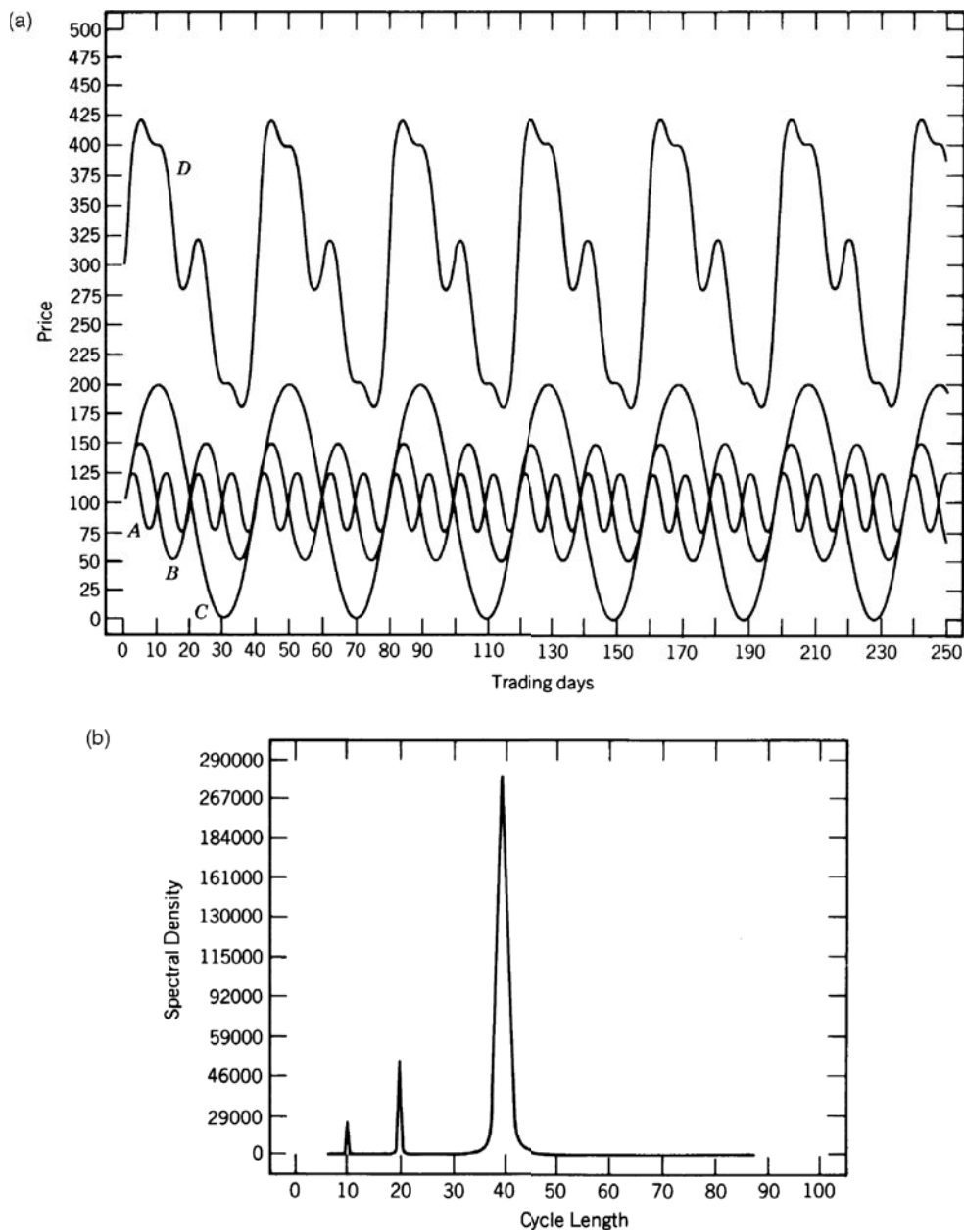
FIGURE 11.12   Spectral density. (a) A compound wave D, formed from three primary waves, A, B, and C. (b) Spectral density of compound wave D.
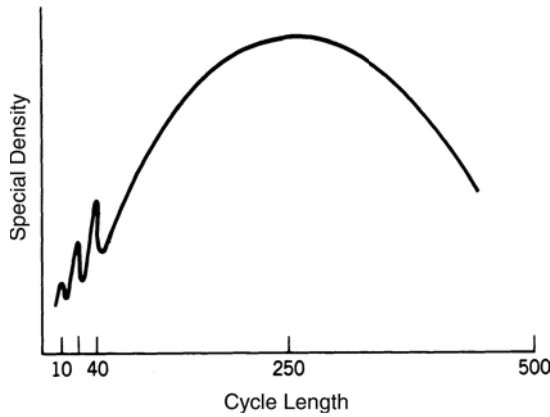
FIGURE 11.13   10-, 20-, and 40-day cycles, within a 250-day seasonal.

Weighting Factors

The most important part of spectral analysis is finding the proper estimators, or weighting factors, for the single-frequency series of cosine waves. When looking for long-term cycles, the trend and seasonal components must be removed because the method of spectral analysis will consider these the dominant characteristics and other cycles may be obscured.

As in the other trigonometric formulas, the basic time-series notation is used, where y_t, $t = 1, 2, \ldots, N$ are the data points and $\hat{y}_t$ will be the resulting estimated points on the spectral analysis. Then

$$\hat{y}_t(\omega) = \frac{1}{\pi}\left(c_0 + 2\sum_{k=1}^{N-1} c_k \cos \omega k \right)$$

where

$$c_k = \sum_{i=1}^{N-k} \frac{(y_t - \overline{y})(y_{t+k} - \overline{y})}{N}$$

Methods of performing spectral analysis vary due to the choice of weighting functions that compensate for the fact that the accuracy of c_k decreases as k increases. The two most popular techniques for handling this problem introduce an estimator λ_k called a *lag window* and a truncation point $M < N$ so that the values of c_k for $M < k < N$ are no longer used and the values of c_k for $k < M$ are weighted by λ_k.

The spectral analysis approximation is then written:

$$\hat{y}_t(\omega) = \frac{1}{\pi}\left(\lambda_0 c_0 + 2\sum_{k=1}^{M} \lambda_k c_k \cos \omega k \right)$$

where λ_k can be either of the following:

Tukey window

$$\lambda_k = \frac{1}{2}\left(1 + \cos\frac{\pi k}{M}\right) \quad k = 0, 1, \ldots, M$$

Parzen window

$$\lambda_k = \begin{cases} 1 - 6\left(\dfrac{k}{M}\right)^2 + 6\left(\dfrac{k}{M}\right)^3 & 0 \le k \le \dfrac{M}{2} \\ 2\left(1 - \dfrac{k}{M}\right)^3 & \dfrac{M}{2} \le k \le M \end{cases}$$

Using a Fast Fourier Transform Program

There are computer programs that apply a *Fast Fourier Transform* to perform a spectral analysis and create a Fourier power spectrum such as the one in Figure 11.12b. Anthony Warren's approach[13] can be found on the Companion Website, *TSM Fast Fourier Transform*, written by John Ehlers in BASIC program code. The program detrends the data and reduces endpoint discontinuity, which can produce large unwanted cycles. This is accomplished by multiplying the data by a bell-shaped *window* and extending the endpoints to give a more definitive structure to the detrended data, without affecting the results (as discussed in the previous section).

A second filter is applied using selected moving averages. The moving average will reduce or eliminate the importance of those cycles, which are equal to or shorter than two times the length of the moving average period, letting the more dominant cycles appear. For example, the use of a 10-day moving average will eliminate cycles of length less than 20 days (frequencies greater than 12.5 per year). Figure 11.14 shows the output of the computer program.

Subsequent works by Warren and Hutson[14] present a computer program to calculate moving average–weighted filters using linear, triangular, and Hanning weights.

Interpreting the Results of the Fourier Power Spectrum

Both Figures 11.12b and 11.14 show a power spectrum resulting from a Fourier transform. Figure 11.12b is an ideal representation, where the cycles stand out with no ambiguity; Figure 11.14 is more realistic, showing both the dominant cycles and a certain amount of variance around those values. In the power spectrum, the cycle power shown

---

[13] Anthony Warren, "A Mini Guide to Fourier Spectrum Analysis," *Technical Analysis of Stocks & Commodities* (January 1983). A very useful series of articles on spectral analysis has been published in *Technical Analysis* beginning in January 1983, authored by both Anthony W. Warren and Jack K. Hutson. Much of the information in this section was drawn from that material.

[14] Anthony Warren and Jack K. Hutson, "Finite Impulse Response Filter," *Technical Analysis of Stocks & Commodities* (May 1983).
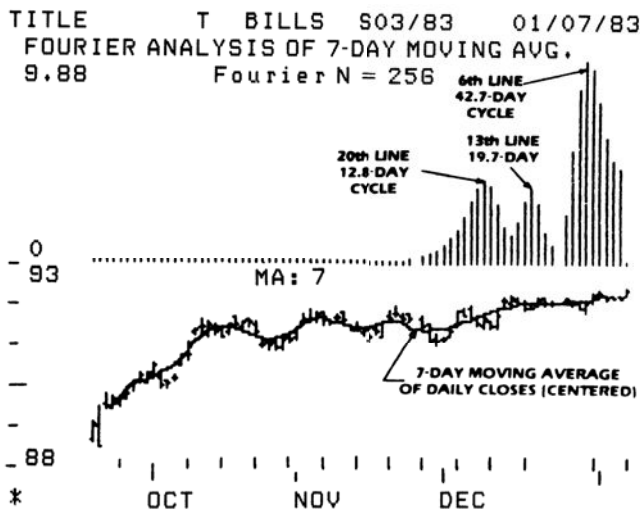
FIGURE 11.14   Output of spectral analysis program.

Source: Jack K. Hutson, "Using Fourier," *Technical Analysis of Stocks & Commodities* (January 1983).

along the *y*-axis is the cycle amplitude squared. In Figure 11.12a cycle *D* peaks at a price of about 425, which yields a spectral density, or *spectral power*, of 180,625 when squared, corresponding roughly to the 40-day cycle in Figure 11.12b.

Using the information from the beginning of this chapter, the frequency is the inverse of the cycle length; therefore, if the cycle length is 40 days, the frequency $F = 360/40 = 9$. The sine wave changes phase at the rate of 9 degrees per day, completing one full cycle every 40 days.

A fast method for observing the possible results is to use weekly rather than daily data. This will be a close approximation for low frequency waves but will be less representative for the high frequencies. Averaging the data points can yield results very similar to the daily analysis.

Using Excel's Fourier Analysis

The Excel spreadsheet program provides a convenient application of Fourier analysis. It is a data analysis *add-in* that is available for most users but must be chosen by the user. This can easily be done by going to *Tools/Add-Ins* and following the instructions for Data Analysis. Once the program has been loaded, it is accessed in the drop-down menu *Tools/ Data Analysis/Fourier Analysis*. The following steps will create the spreadsheet shown in Figure 11.15.[15]

---

[15]The author gratefully acknowledges the help of John Ehlers in the interpretation of Excel's Fourier Analysis.

1. Load your data into a worksheet. Monthly data are best for finding longer-term cycles, ranging from three months to a few years. In Figure 11.15 the data (date and price) are shown in the first two columns.

2. The Fourier analysis program will only take the amount of data equal to a power of 2, that is, 2, 4, 8, 16, . . . , 1024, 2048, 4096. It will not accept more than 4096 data values. Using monthly data keeps the quantity well within those constraints; however, you may need to use only the most recent 256 or 512 data points.

3. Remove the trend by taking the first differences, or the returns, of the data. The first differences are $C_t - C_{t-1}$, carried down in the column marked "Diff." Differences are good if the data is stationary, but returns are best if prices continue higher; otherwise, more recent data, at higher prices, will overwhelm the results. In the example, both differences and returns are compared.

4. Go to *Tools/Data Analysis/Fourier Analysis* and enter the data points from the column "Price" as the input, and the corresponding cells from the first empty column, E. Column E will become the output, "Fourier result using returns."

5. Note that most of the numbers in column E are *imaginary*. The term *imaginary* is a mathematical term that refers to the result of taking the root of a negative number. To create a usable value, apply the function *IMABS* to each value in column E and store it in column F.

Only the first 12 rows of the spreadsheet are shown in Figure 11.15. The sixth column (IMABS) will hold the significant monthly cycles 1 through 12. The larger numbers indicate a dominance of that monthly cycle. In Figure 11.16, the cycle strength for both the returns and the differences is shown as a bar chart, which is intended to represent the spectral density. The pattern can be interpreted as the seasonality for Southwest Airlines. Both the returns (the bars) and differences (the line) show about the same results, although the magnitude of the values is very different. The chart shows stock prices rising through the summer, and again then in November and December. Lows are in January, May, and October. October may be distorted due to the 9/11 terrorist attack; otherwise,

| Date	Price	Diff	Returns	Fourier Results Using Returns	IMABS
8/3/1998	9.46	0	0	0.501146645850281	0.501
8/31/1998	7.78	−1.68	−0.17759	0.70487176607942−0.166330457208832i	0.724
9/30/1998	8.84	1.06	0.136247	1.02042917861071+0.661593634349334i	1.216
10/30/1998	9.22	0.38	0.042986	0.146079614330543+0.84883419259415i	0.861
11/30/1998	9.47	0.25	0.027115	2.2133471150028E-002+0.10491904750043i	0.107
12/31/1998	9.91	0.44	0.046463	−0.863452208950358−1.00760775938179i	1.327
1/29/1999	11.77	1.86	0.187689	0.259591537370051−1.00408817698152i	1.037
2/26/1999	13.16	1.39	0.118097	1.47274269096796+9.00145861328777E−002i	1.475
3/31/1999	13.22	0.06	0.004559	2.02382065563341+0.226798030523937i	2.036
4/30/1999	14.23	1.01	0.076399	−0.655478248460181+2.87122220450429E−002i	0.656
5/28/1999	14.01	−0.22	−0.01546	−0.761473092874587−0.719913003039197i	1.048
6/30/1999	13.6	−0.41	−0.02926	0.496238674586104−1.22248411855721i	1.319

FIGURE 11.15   Initial rows of Excel spreadsheet showing input and output from Fourier analysis for Southwest Airlines (LUV), August 1998 through April 2011.
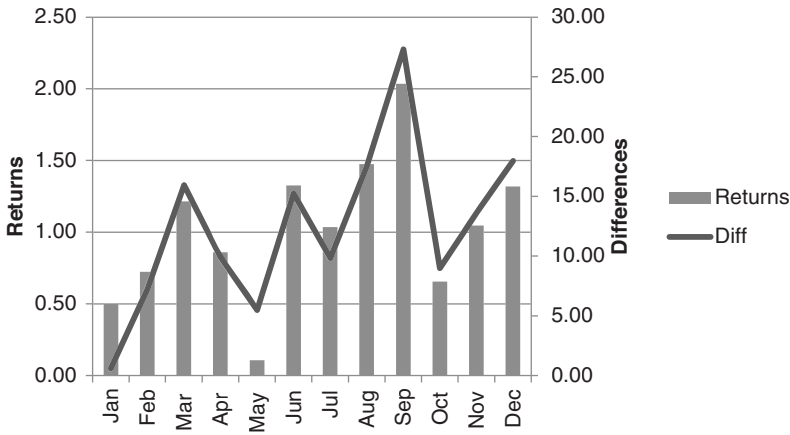
FIGURE 11.16 Results of Excel Fourier analysis for Southwest Airlines (LUV), August 1998 through April 2011.

this may be a reasonable representation, not of when people fly, but of when they buy tickets, or when investors believe that Southwest income will reflect revenues.

In order to have a basis for comparison, another Fourier analysis was performed on cash corn prices from 1990 through April 2011. Figure 11.17 also shows that both results, using differences and returns, are very similar; however, returns, which are percentages, will adjust for some of the distortion due to much larger price changes in recent years. While the scale of the numbers is different from the seasonal study in Chapter 10, the cyclic low in September corresponds to selling pressure during harvest, and the three stronger summer months confirm our intuitive conclusion that the growing season breeds uncertainty.
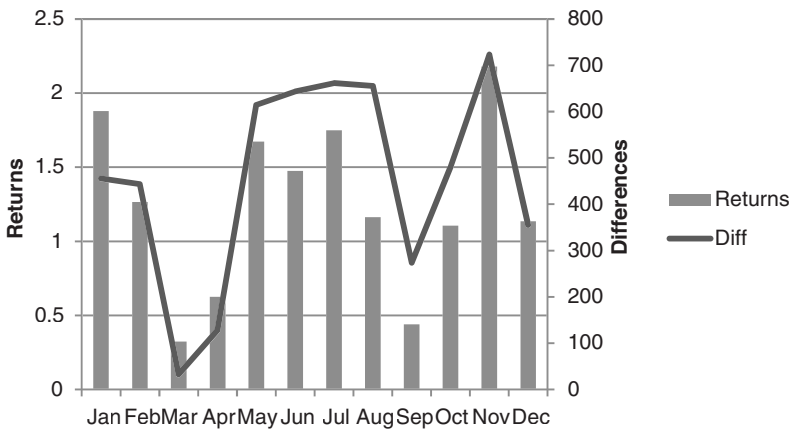
FIGURE 11.17 Cash corn cycles using Excel's Fourier analysis based on returns and price differences, 1990 through April 2011.

MAXIMUM ENTROPY

Maximum Entropy Spectral Analysis (*MESA*) is a technique created by John Ehlers[16] that filters noise (*entropy*) from a time series and exposes the useful cycles.[17] It provides a very practical alternative to Fourier analysis and makes it possible to find cycles using a very small amount of data by making the data appear more stationary. The use of Fourier transforms requires at least 256 data points and a minimum of 16 consistent cycles of 16 bars. That would eliminate the possibility of uncovering cycles for the short-term trader.

Ehlers describes the existence of short-term cycles as a natural phenomenon.[18] It is part of the process that causes rivers to meander back and forth as water seeks to flow in a straight line, or a drunkard who walks through an alley bumping against the walls but moving steadily forward. From these patterns, useful cycles can be found about 20% of the time. The presence of a useful cycle, or the lack of one, can be a valuable piece of information for a trader.

Using the Phase Angle

In an ideal situation, where the market cycle can be shown as a pure sine wave, the phase angle constantly increases throughout the cycle, beginning at 0° and ending at 360° (also equal to 0°). The phase angle then drops to zero when the new cycle begins and increases again at a constant rate until it ends at 360°. This repeated pattern forms a *sawtooth* chart shown in Figure 11.18 (further description of the phase angle can be found in the upcoming section on the Hilbert Transform). Although the cycle goes from peak to valley, the phase angle moves constantly in one direction.

In the practical analysis of short-term cycles, Ehlers compresses tick data into bars of equal numbers of ticks and then examines the phase for uniformity. Once found, the uniform phase, which appears as a sawtooth chart, will become erratic as the short-term cycle begins to break down, marking the end of the current market event. Ehlers creates an *instantaneous trendline* after removing the dominant cycle, and uses it together with a smoothed price trendline to generate trading signals. A trend exists if the smoothed price line has not crossed the instantaneous trendline within the last half-dominant cycle. The mathematics of MESA is extensive but Ehlers makes it remarkably clear, along with simple summaries of important points, in the second edition of *MESA and Trading Market Cycles* and in the newer *Cybernetic Analysis of Stocks and Futures*.

---

[16]John F. Ehlers, *MESA and Trading Market Cycles*, second edition (New York: John Wiley & Sons, 2001); also in *Cybernetic Analysis for Stocks and Futures* (Hoboken, NJ: John Wiley & Sons, 2004).

[17]Anthony Warren, "An Introduction to Maximum Entropy Method (MEM)," *Technical Analysis of Stocks & Commodities* (February 1984). See the bibliography for other articles on this topic.

[18]John F. Ehlers, "How to Use Maximum Entropy," *Technical Analysis of Stocks & Commodities* (November 1987).
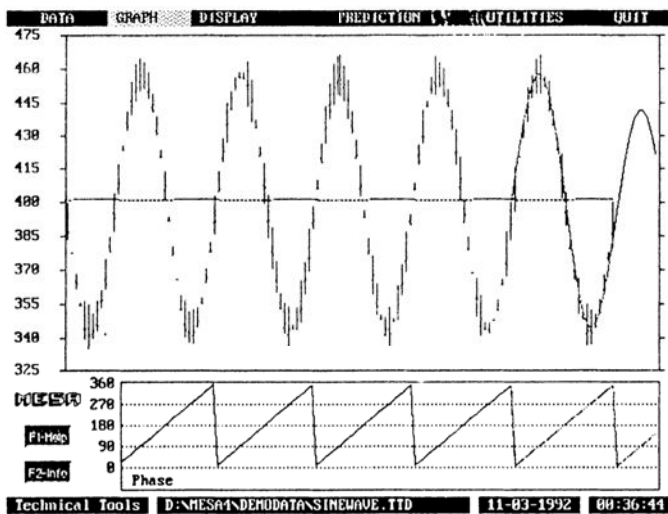
FIGURE 11.18   The phase angle forms a sawtooth pattern.

Source: John Ehlers, "Cycle Analysis and Intraday Trading,"
Technical Analysis of Stocks & Commodities 11, no. 4 (February
1993). © 1993 Technical Analysis, Inc. Used with permission.

Ehlers' Lateral Shift in Thinking

At first glance the use of only a small amount of data needed by MESA seems to contradict the basic rules of statistics, which demand that results be based on as much data as possible in order to be reliable. But Ehlers, who has been the dominant influence in cycles since 1990, is too knowledgeable to have made such a simple mistake. His book *MESA and Trading Market Cycles*, which focuses on the use of short-term cycles based on short sample time periods, used this very attribute turned inside out when applied to cycles.

Ehlers' objective is to find very short-term cycles. By definition, these cycles must be the result of human behavior, rather than based on market economics, because fundamentals are not usually relevant to periods of only a few days, and are not likely to have a regular pattern when they make a rare appearance. If very short-term cycles exist, they will not continue for long periods, and you must recognize them quickly if they are to be useful; therefore, short-term cycles are found by analyzing only a small amount of recent data.

Then why bother to find a short-term cycle based on a small amount of data, if it is not statistically dependable? In a lateral shift, Ehlers uses the existence of a short-term cycle to tell if prices are in a sideways pattern or trending. If a short-term cycle exists, then the market cannot be trending. Ehlers has no interest in trading the cycle, which is surprising for a cycle expert, but prefers the dependability of the trend. He has, instead, attempted to solve one of the most difficult problems facing the analyst, trying to distinguish between a trending and sideways market. *If a short-term cycle exists, then the market is not trending.* Ehlers develops this method throughout his book.

Finding the Cycle Using the Hilbert Transform

Ehlers is able to recognize the cyclic component of price movement using very little data as contrasted with the traditional regression methods. In one technique, using the *Hilbert Transform*,[19] only a small part of one cycle is needed to form a picture of the entire process, as little as 4 bars. This allows the cycle to be shown as an indicator with only a modest amount of lag.

The Hilbert Transform is based on the separation of the cycle phase, represented by a *phasor*, into two components, the *Quadrature* and *InPhase*, shown in Figure 11.19. The left circle with a single arrow points to the current position of the cycle based on the phasor being straight up (270°) at the cycle peaks and straight down (90°) at the cycle valleys. The cycle begins when the phasor is pointing to the right (0°). The right circle separates the phasor into its horizontal and vertical components, InPhase and Quadrature, respectively. The *phase angle*, shown as θ (theta), is the arctangent of the ratio of the Quadrature and InPhase components. Ehlers reduces the equations for the Hilbert Transform to:

Quadrature
$$Q = 0.0962 \times price_t + 0.5769 \times price_{t-2} - 0.5769 \times price_{t-4} - 0.0962 \times price_{t-6}$$
InPhase
$$I = price_{t-3}$$

These equations make it possible to write simple program indicators (*TSM Hilbert Transformation* and *TSM Hilbert Transformation V2*, available on the Companion Website) to plot the results of the Hilbert Transform for any data series. Although there are some penalties for truncating the Hilbert Transform, which is an infinite series, those penalties should not affect the use of this method for trading.
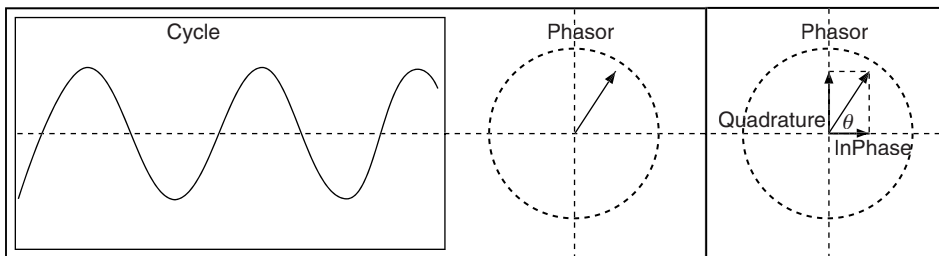
FIGURE 11.19 A cycle with the phasor and phase angle.
Source: Adapted from John F. Ehlers, *Rocket Science for Traders* (New York: John Wiley & Sons, 2001).

---

[19] John F. Ehlers, *Rocket Science for Traders* (New York: John Wiley & Sons, 2001), Chapter 6.
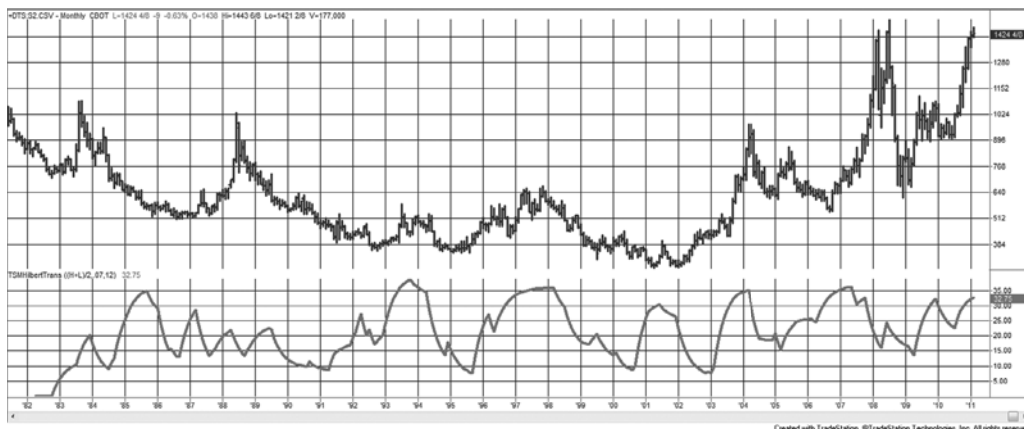
FIGURE 11.20 Back-adjusted soybean futures from 1982 with the Hilbert Transform in the lower panel.

Applying the Hilbert Transform to soybean monthly prices gives the results shown in Figure 11.20. The continuous, back-adjusted series is shown at the top of the chart. Continuous data is preferable in this case to avoid any odd jumps in prices when one contract is rolled to another at expiration.

The bottom panel in Figure 11.20 is the result of first detrending the monthly data (part of the calculation process), then applying the Hilbert Transform with $alpha = 0.07$. The Hilbert Transform creates sharp peaks and valleys at the points where the cycle is expected to be at maximum and minimum value. There is usually one peak and one valley each year, and the minimum level often coincides with late summer or early fall, when the harvest yield is known and there is usually a surplus (or anticipated surplus) of soybeans. Sharp peaks are particularly good for mean reversion trading compared to indicators that reach maximum or minimum values and stay there while prices continue in the same direction. However, as we saw when comparing the results of cash and back-adjusted futures, these extremes will vary due to forward discounting of prices. Peaks tend to be earlier in the year using futures data; nevertheless, futures data are the only practical means of trading. In addition, geopolitical events have been much more influential in the past 10 years, and these are more likely to affect the high prices than the lows. A more thorough discussion can be found in Chapter 10. The same method can be applied to weekly data. If the obvious major cycle is yearly, then a 52-period average can be used to detrend the data.

The Hilbert Transform indicator does a very good job of locating relative peaks and the highest and lowest values of the indicator could be used for sell and buy signals. However, the chart of the indicator does show characteristics similar to momentum indicators; that is, larger peaks in the indicator follow periods of low volatility in prices because the subsequent peaks will be seen as *relative highs*. While the peak to valley might be an ideal trade, the net returns will vary due to individual market volatility.

The Fisher Transform

It is well known that prices are not normally distributed; that is, the distribution of price changes, or price minus a trendline (detrended prices), does not appear to be a bell-shaped, symmetric curve. We have already discussed some of the idiosyncrasies of price movement in earlier chapters, including the fat tail of trend-following performance, or the increase in volatility with price. The distribution of prices is called a *probability density function* (PDF), and the normal, bell-shaped curve is a *Gaussian PDF*.

The way in which prices move between two bands is very similar to the probability density function of a sine wave,[20] which spends more time in the vicinity of the peaks and valleys (where it changes direction) than in the middle (where it moves the fastest). Figure 11.21a shows two cycles of a sine wave with the PDF to the right (Figure 11.21b). Although the PDF is normally shown with the phasor angle along the bottom, as in (Figure 11.21c), this chart is drawn to represent a typical frequency
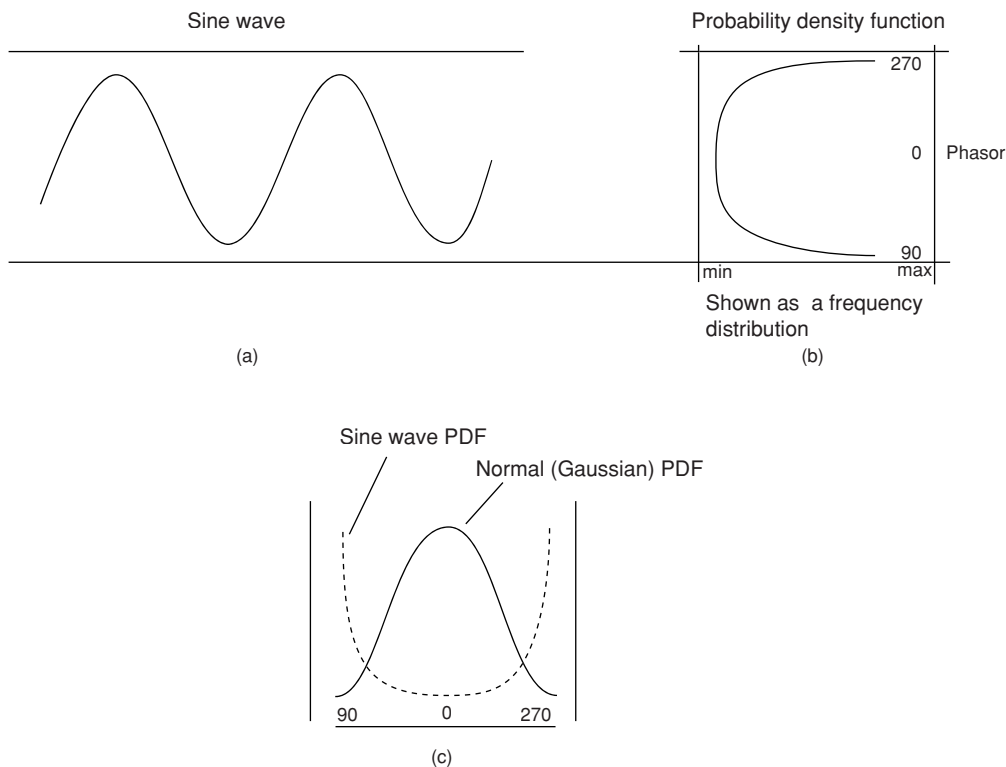
FIGURE 11.21 Probability Density Function (PDF) of a sine wave.

---

[20]This section is adapted from John Ehlers, *Cybernetic Analysis for Stocks & Commodities* (Hoboken, NJ: John Wiley & Sons, 2004), Chapter 1.

distribution. The peaks of the sine wave occur when the phasor angle is 270° and the lowest points when the angle is 90°. The frequency of the peaks (the top of the chart) and valleys (the bottom of the chart) are much greater than the frequency of the other angles, especially 0° and 180°.

The same PDF can be seen in price movement if we form a channel around the prices and measure the relative position of prices within that channel. The channel high (*MaxH*) is simply the maximum price during period p, and the channel low (*MinL*) is the minimum price during the same period p. The value used in the distribution is calculated

$$X_t = 0.5 \times 2 \times ((P_t - MinL)/(MaxH - MinL) - 0.5) + 0.5 \times x_{t-1}$$

The Fisher Transform takes this distribution and changes it to one that is approximately Gaussian with the following formula:

$$y = 0.5 \times \ln\left(\frac{1+x}{1-x}\right)$$

where x = the input
y = the output
ln = the natural log

The result of applying the Fischer Transform to AMR monthly data can be seen in the second panel of Figure 11.22. For comparison, the Hilbert Transformation is in the third panel. In this case, the period for the calculation of the bands was 12 months, in order to allow a better comparison with other methods. An alternate choice would be 3 months to correspond to calendar quarters and earnings reports. Values for the Fisher Transform range from +1.0 to −1.0. The peaks of the Fisher Transformation are remarkably in line with the price peaks, and show very little lag compared to the Hilbert Transformation,
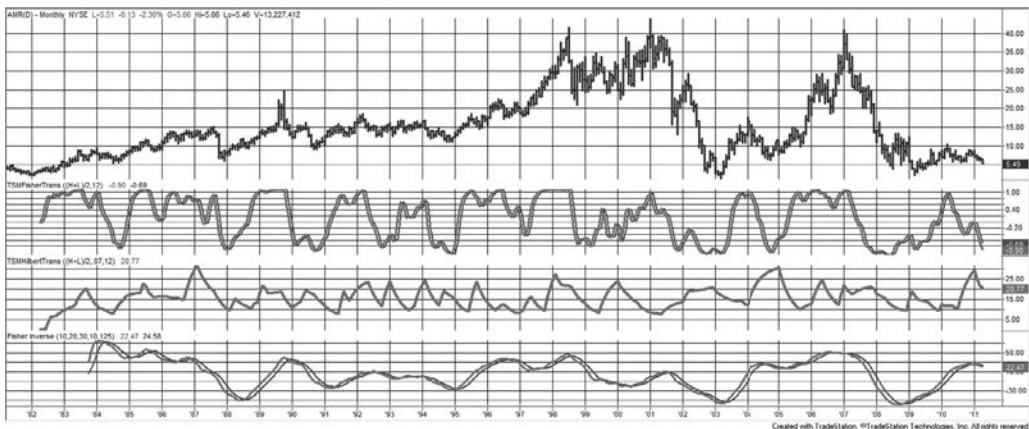
FIGURE 11.22 Monthly AMR (top) prices from 1982 with the Fisher Transform (2nd panel), Hilbert Transform (3rd panel), and Inverse Fisher Transform (lower panel)

although they occasionally peak out early and hold that level until prices reverse. The bottoms are also good, although there is an occasional lag. The Fisher Transform produces clearer, sharper turning points than a typical momentum-class indicator. A *trigger* is also included that corresponds to the MACD *signal line*. A sell signal occurs when the Fisher Transform value crosses the trigger moving lower. Experience shows that the best signals are those occurring just after an extreme high or low value and not after a turning point where the value is near zero, similar to MACD rules. Programs for creating the Hilbert, Fisher, and Inverse Fisher transforms are *TSM Hilbert Transform*, *TSM Fisher Transform*, and *TSM Fisher Inverse*, which can be found on the Companion Website.

Inverse Fisher Transform

One more variation produces a very credible momentum indicator, the *Inverse Fisher Transform.*[21] While the Fisher Transform uses the price to solve for the distribution, the Inverse Fisher Transform does the opposite,

$$x = \frac{e^{2y} - 1}{e^{2y} + 1}$$

However, implementation of this is wrapped around the RSI indicator as follows (shown in TradeStation code):

```
           series = close - close[mom];
        x1 = 0.1*(RSI(series,RSIper) - 50);
             x2 = WAVERAGE(x1,waper);
  invfisher = 100*(EXPVALUE(2*x2) - 1.0)/(EXPVALUE(2*x2) + 1);
```

using the input values, *mom* = 10, *RSIper* = 20, and *waper* = 30. Note that *RSI* and *WAVERAGE* are the functions for the RSI and the linearly weighted average.

This process gives a bipolar distribution, where results are most likely to cluster near the extremes, +1 and −1. Based on monthly data of AMR, the results of the Inverse Fisher Transform are shown in the bottom panel of Figure 11.22.

CYCLE CHANNEL INDEX

A trend-following system that operates in a market with a well-defined cyclic pattern should have specific qualities that do not exist in a basic smoothing model. In order to confirm the cyclic turning points, which do not often occur precisely where they are expected, a simple moving average should be used, rather than an exponentially smoothed one. Exponential smoothing always includes some residual effect of older data, while the moving average uses a fixed period that accommodates the characteristics of a cycle.

---

[21] See www.mesasoftware.com/papers/ and follow the link.

The cyclic turning point will use part of the data that represents about ¼ of the period, combined with a measure of the relative noise in the series which may obscure the turn.

These features have been combined by Lambert[22] into a *Commodity Cycle Index* (CCI), which is calculated as:

$$\mathrm{CCI}_t = \frac{x_t - \bar{\bar{x}}_t}{0.015 \times \mathrm{MD}_t}$$

where $x_t = (H_t + L_t + C_t)/3$ is the average of the daily high, low, and close

$\bar{x}_t = \displaystyle\sum_{i=1-N+1}^{t} x_i,$ the moving average over the past N days

$MD_t = \displaystyle\sum_{i=1-N+1}^{t} \left| x_i - \bar{x} \right|,$ the mean deviation over the past N days

$N =$ the number of days selected (less than ¼ cycle)

Because all terms are divided by N, that value has been omitted in the final formula. In the CCI calculations, the use of $0.015 \times \mathrm{MD}$ as a divisor scales the result so that 70% to 80% of the values fall within a +100 to −100 channel. The rules for using the CCI state that a value greater than +100 indicates a cyclic turn upward; a value lower than −100 defines a turn downward. Improvements in timing rest in the selection of N as short as possible but with a mean-deviation calculation that is a consistent representation of the noise. The CCI concept of identifying cyclic turns is good because it accounts for the substantial latitude in the variance of peaks and valleys, even with regular cycles.

SHORT CYCLE INDICATOR

In an excellent article,[23] Francisco Lorca-Susino presents the *Short Cycle Indicator*. This method is expected to correct some of the difficulty in financial time series, which are said to have a *long memory*, obscuring some of the patterns. It is applied to intraday bars and is best interpreted over multiple time frames.

The formula is based on the squared difference of two exponential moving averages, and the relationship of those trendlines with the highest low and lowest high of the slower period, a form of stochastic indicator. The term $(XF - XS)/XF$ addresses

---

[22] Donald R. Lambert, "Cycle Channel Index," *Commodities* (1980), reprinted in *Technical Analysis of Stocks & Commodities*.

[23] Francisco J. Lorca [Susino], "Exploiting Stock Market Cycles," *Futures* (April 2009). The author appreciates the help of Mr. Lorca-Susino in this section. He can be reached at franlorcasusino@gmail.com.

the convergence of the two trends. The combination of squaring and the stochastic is intended to smooth while still retaining the most important aspects of sensitivity.

$Slow = 20$

$Fast = 8$

XF = fast exponential smoothing *(Fast)*

XS = slow exponential smoothing *(Slow)*

LH = Lowest*(High,Slow)*

HL = Highest*(Low,Slow)*

$SF_1 = 1,000,000$

$SF_2 = 100$

$SF_3 = 1000$

$$Lorca1 = \left(\frac{(XS - LH)^2 + (XS - HL)^2}{XS} \right) \times \frac{XF - XS}{XF} \times SF_1$$

$$Lorca2 = \frac{(XS - LH)^2 + (XS - HL)^2}{XS} \times SF_2 \times \frac{\dfrac{XF - XS}{XF} \times SF_3}{C^2} \times SF_1$$

As seen in Figure 11.23, the indicator tends to stay above or below the trigger line but reacts to changing volatility, recognized as divergence of the trendlines and the price extremes. These usually occur before prices change direction. It is interesting to

FIGURE 11.23 The Short Cycle Indicator applied to the *emini* S&P 30- and 60-minute bars, May 2011.

note the multiple divergence signals that occur as prices rally on the first part of the chart, and that the indicator posts its lows in advance of the lows on April 5.

To generate a *sell* signal, Lorca looks for volatility to expand in a particular pattern. One of the three conditions must be satisfied

1. $Lorca1_t > 0$ and $Lorca1_t < Lorca1_{t-12}$ and $Lorca1_t < Lorca1_{t-14}$ and $Lorca1_t < Lorca1_{t-16}$ and $H_t - L_t > H_{t-1} - L_{t-1}$ and $H_t = \text{Highest}(H,12)$
2. $Lorca1_t > 0$ and $H_t = \text{Highest}(H,16)$ and $Lorca1_t = \text{Lowest}(Lorca1,32)_{t-1}$
3. $Lorca1_t > 0$ and $H_t = \text{Highest}(H,16)$ and $Lorca1_{t-2} = \text{Lowest}(Lorca1,32)_{t-1}$

In case 1, the indicator is rising but it is lower than the interval from 12 to 14 bars ago. In addition, the trading range of the current bar must be greater than the range of the previous bar and confirmed by a new 12-bar high, both measures of expanding volatility.

Cases 2 and 3 are simpler, requiring that the indicator is rising and that this is a 16-bar high. Then a *sell* signal occurs if either the current indicator value is the lowest of the past 32 bars or the indicator 2 bars ago is the lowest of the past 32 bars.

PHASING

One of the most interesting applications of the cyclic element of a time series is presented by J. M. Hurst in *The Profit Magic of Stock Transaction Timing.*[24] He uses *phasing*, the synchronization of a moving average, to represent cycles. This section highlights some of the concepts and presents a simplified example of the method. It is already known that to isolate the cycle from the other elements, the trending and seasonal factors should be subtracted, reducing the resulting series to its cyclic and remaining noise parts. In many cases, the seasonal and cyclic components are similar but the trend is unique. Hurst treats the cyclic component as the dominant component of price movement and uses a moving average to identify the combined trend-cycle.

The system can be visualized as measuring the oscillation about a straight-line approximation of the trend (a best-fit centered line), anticipating equal moves above and below. Prices have many long- and short-term trends, depending on the interval of analysis. Because this technique was originally applied to stocks, most of the examples used by Hurst are long-term trends expressed in weeks. For futures the same technique could be used with continuous back-adjusted data.

As a simple example of the concept, choose a moving average of medium length for the trending component. The *full-span* moving average period may be selected by averaging the distance between the tops on a price chart (a rough measure of the cycle). The *half-span* moving average is then equal to half the days used in the full-span average.

---

[24]J. M. Hurst, *The Profit Magic of Stock Transaction Timing* (Upper Saddle River, NJ: Prentice Hall, 1970), reprinted in trade paper by Traders Press, 2000.
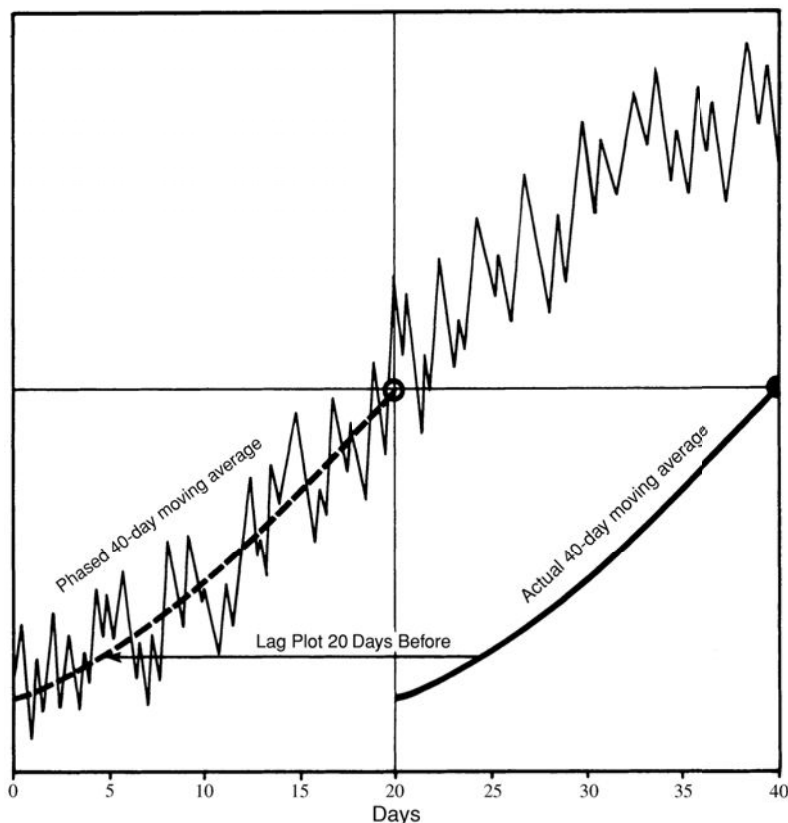
FIGURE 11.24   Hurst's phasing.

The problem with using moving averages is that they always lag. A 40-day moving average is considered to be 20 days behind the price movement. The current average is normally plotted under the most recent price, although it actually represents the average of the calculation period and could be lagged by one-half the period. Hurst's method applies a process called *phasing*, which aligns the tops and bottoms of the moving average with the corresponding tops and bottoms of the price movement. To phase the full- and half-span moving averages, lag each plot by half the days in the average; this causes the curve to overlay the prices (Figure 11.24). Then project the phased full- and half-span moving averages until they cross. A line or curve connecting two or more of the most recent intersections will be the major trendline. The more points used, the more complicated the regression formula for calculating the trend; Chapter 6 discusses a variety of linear and nonlinear techniques for finding the best fit for these intersections, although a basic least-squares solution might be best. Once the trendline is calculated, it is projected as the center of the next price cycle.

With the trend identified and projected, the next step is to reflect the cycle about the trend. When the phased half-span average turns down at point *A* (Figure 11.25), measure
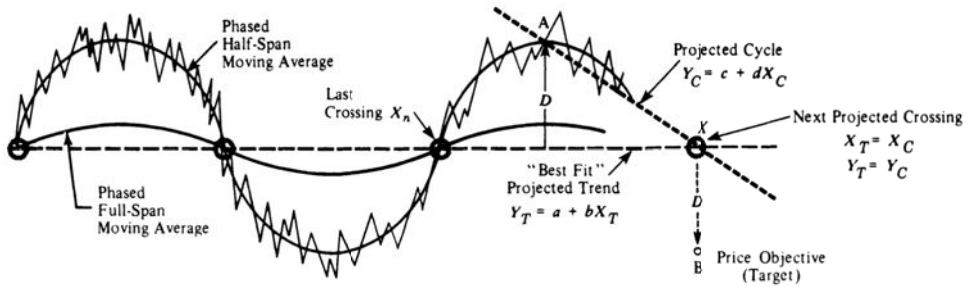
FIGURE 11.25 Finding the target price.

the greatest distance D of the actual prices above the projected trendline. The system then anticipates that prices will cross the trendline at point X and decline an equal distance D below the projected centered trendline. Once the projected crossing becomes an actual crossing, the distance D can be measured and the exact price objective specified. The rules for using this technique are:

1. Calculate the full-span moving average for the selected calculation period; lag the plot by half the days. If the full-span moving average uses F days, the value of the average is calculated at $t - F/2$, where t is the current day. Call this phased point PH_t.
2. The half-span moving average is calculated for H days and plotted at $t - H/2 + PH_t$.
3. Record the points where the two phased averages PH_i and PF_i cross and call these points $X_n, X_{n-1}, \ldots$.
4. Find the trend by performing a linear regression on the crossing points $X_n, X_{n-1}, \ldots$. For the straight line, $Y_T = a + bX_T$.
5. Record the highest (or lowest) values of the price since the last crossing, X_n.
6. Calculate the projection of the half-span by creating a straight line from the highest (or lowest) half-span value since the last crossing A to the last calculated half-span value. This equation will be $Y_c = c + dX_c$.
7. Find the point at which the projected trendline crosses the projected cyclic line by setting the equations equal to one another and solving for X and Y. At the point of crossing $(X_T, Y_T) = (X_c, Y_c)$, giving two equations in two unknowns, which is easily solvable (X is time in days; Y is price).
8. If the half-span is moving down, the maximum price reached since the last crossing is subtracted from the Y coordinate of the projected crossing to get D. This distance D is subtracted again from the Y coordinate to determine the price objective, equidistant below the best-fit line. If the half-span is moving up, the price objective uses the minimum price and reflects the distance above the projected crossing. It should be noted that this calculation of distance is simplified because the trend is established by a straight line; for nonlinear fits, the measurement of D will be more complicated.

9. Recalculate the moving averages (Step 1), the half-span projection (Step 6), the projected crossing (Step 7), and price objective (Step 8) each day until the actual crossing occurs. At that time D is fixed.

10. Follow the trading rules:

    a. *Enter a new long position* when the half-span moving average turns up; cover any existing short positions regardless of the price objective.

    b. *Enter a new short position* when the half-span moving average turns down; close out any long positions.

    c. *Close out both long and short positions* if the price objective is reached. An error factor of 10% of the height of the full cycle (lowest to highest point) should be allowed. Therefore, the price objective should be closer by 10%.

Hurst's approach is a good example of a complex problem solved using elementary mathematics. There are many techniques for determining trends and a number of seasonally oriented systems, but a cyclic approach is rare. Hurst's explanation is more complete and the interpretation presented in this section should be considered only as a reasonable approximation.

Volume, Open Interest, and Breadth

P atterns in volume have always been tied closely to chart analysis of both the stock and futures markets. Volume is a valuable piece of information that is not often used, and one of the few items, other than price, that is traditionally considered valid data for the technician. Nevertheless, there has been little research published that relates volume to futures markets; its popular use has assumed the same conclusions as in stock market analysis.

The stock and futures markets have two other measures of participation that are related. In equities, the large number shares being traded allow for measurements of *breadth.* In the same way that the stock index has become a popular measure of overall market trend, the *breadth of the market* is the total number of stocks that have risen or fallen during a specific period. When you have the ability to view the bigger picture of market movement, breadth seems to be a natural adjunct to the index. In addition, tallies of *new highs* and *new lows* may add value to a trading decision. This chapter will look at different way these statistics can complement other indicators.

In futures, *open interest* is the measurement of those participants with outstanding positions; it is the netting out of all open positions (longs and short sales) in any one market or delivery month, and gives an understanding of the depth of participation and antici-pated volume. A market that trades only 10,000 contracts per day but has an open interest of 250,000 is telling the trader that there are many participants who will enter the market when the price is right. These are most likely to be commercial traders, using the futures markets for hedging. Unlike the stock market, which can only trade a fixed number of out-standing shares, the futures market can add net contracts for every new buyer and short seller and reduce the *interest* in the market when both buyers and sellers liquidate.

A SPECIAL CASE FOR FUTURES VOLUME

In futures markets, where you trade contracts for delivery in specific months, the volume of each contract is available, along with the total volume of all contracts traded for that

market. Spread transactions within that market, for example, buying the March delivery and selling June, are not included in the reported volume, a point that is inconsistent with stocks, which report both sides of a high frequency trade. In futures, volume data and open interest are officially posted one day late, but estimates are available for many markets during the day on services such as Bloomberg. Total volume of crude oil is estimated every hour; it is released to news services and is available on the Internet.

Individual contract volume is important to determine which delivery month is most active. Traders find that the best executions are where there is greatest liquidity. Analysts, however, have a difficult time assessing trends in volume because there is a natural increase in volume as a contract moves from deferred to delivery, from second month out to the nearby, and traders most often roll their positions to the closest delivery month. There is a corresponding decline in volume as a contract approaches its delivery date. Because spread volume is not posted, movement between delivery months does not affect the volume patterns.

Each futures market has its unique pattern of volume for individual contracts. Some, such as the interest rates, shift abruptly on the last day of the month prior to delivery, because the exchange raises margins dramatically for all traders holding positions in the delivery month. Currencies are very different and tend to trade actively in the nearest month up to one or two days before that contract goes *off the board*. While volume increases slightly in the next deferred contract, anyone trading sizable positions will need to stay with the nearby contract to just before the end.

Other than for determining which contract to trade, and perhaps the size of an order that the market can absorb, an analysis of volume as discussed in this chapter will use total volume (the aggregate of all contracts) in order to have a series that does not suffer the patterns of increasing and decreasing participation based on the coming and going of individual delivery months. When traders roll from the nearby to the next deferred contract, the transactions are usually executed as a spread, and those trades are not included in the volume figures. Because positions are closed out in one contract and opened in another there is no change in the open interest.

The stock market equivalent to using total volume would be to add the volume for all stocks in the same sector or industrial group. This would help to smooth over those periods when the volume is concentrated in a few stocks following a news or earnings release, only to switch the next week when other stocks in the same sector are noticed.

Tick Volume

The popularity of quote machines and fast trading requires a measurement of volume that can be used immediately to make decisions. Lacking a service such as Bloomberg with continually updated volume, *tick volume* can be substituted. Tick volume is the number of recorded price changes, regardless of volume or the size of the price change that occurs during any time interval. Tick volume relates directly to actual volume because, as the market becomes more active, prices move back and forth from bid to asked more often. If only two trades occur in a 5-minute period, then the market is not liquid. From

an analytic view, tick volume gives a reasonable approximation of true volume and can be used as a substitute. From a practical view, it is the only choice. Higher-than-normal tick volume at the beginning of the day implies higher volume throughout the day. Tick volume patterns are discussed later in this chapter.

VARIATIONS FROM THE NORMAL PATTERNS

The *W* Intraday Pattern

One note of warning when using intraday volume or volume indicators to confirm price direction: The patterns in volume have a dominant *W* pattern throughout the day. They begin high, drop quickly, increase modestly near midday, fall again, then increase significantly toward the close of trading. To decide that a buy signal is more important near the end of the trading day because volume was rising is not correct; volume is always higher at the beginning and end of the day. You may conclude, in general, that the beginning and end of the day, or higher-than-average volume produces more reliable trading signals; however, an intraday volume confirmation must be compared against the normal volume for that time of day.

Open interest and market breadth have seasonal patterns. In agricultural markets, farmers hedge in larger numbers during the growing season than in the winter, raising the open interest. In stocks, there is a lot of activity associated with the end-of-year positioning for tax purposes and traditional rallies during holiday seasons. Volume is low during the summer when many investors take their vacations. None of these changes indicate that something special is occurring. These variations are discussed in detail later in this chapter.

Variance in Volume

Nearly all volume analysis uses smoothed or averaged data because volume can vary substantially from one day to the next. As an example, Table 12.1 shows the breakdown of Microsoft (MSFT) volume for the five years from June 1998 through May 2003. While the volume was higher during the price peak of January 2000, the overall numbers show consistency at an average of 71 million shares per day.

Table 12.1 also shows the standard deviation of the daily volume changes, which is about 38% of the average volume. One standard deviation of this magnitude says that 68% of the days will show *changes* in daily volume ranging from 41 million to 101 million shares. Looking at it in reverse, 32% of the days will show changes in volume less than or greater than 41 million and 101 million shares. Compare that to the price of Microsoft during the same period, which ranged from $20 to $60 but had 1 standard deviation of $0.957—that is, the daily change in price was less than $1, or about 2.9% of the $33.38 average price. Keep in mind that this variation is the basis for most of the volume indicators and strategies.

TABLE 12.1 Microsoft Daily Volume by Year from June 1998 through May 2003

Period	Average	StDev
Jun98–May99	62,861,035	21,511,856
Jun99–May00	64,512,077	34,741,178
Jun00–May01	83,915,101	35,621,716
Jun01–May02	64,638,359	21,125,063
Jun02–May03	79,584,139	22,129,923
Jun98–May03	*71,105,478*	*27,832,841*

Volume Spikes

Madness is the exception in individuals but the rule in groups.

—Nietzsche

A *volume spike* is a single day on which the volume was much higher than the previous day—at least twice as high, perhaps three or four times. A volume spike is a warning that something happened, most likely the result of a surprising news release or new economic data. It could be the crescendo of a few days of rising volume associated with sharply rising or falling prices.

A volume spike is a clear, positive action by investors. It implies that a very large number of investors, perhaps even the general public, all hold the same opinion on the direction of the market and feel compelled to act on that opinion at the same time. It is the result of mass behavior discussed by Mackay in his famous book, *Extraordinary Popular Delusions and the Madness of Crowds*. A volume spike means that everyone has jumped into the boat at the same time. Traditional interpretation of a volume spike is that it indicates the end of a price move, that is, the boat sinks.

AOL shows three good examples of volume spikes in Figure 12.1. The highest volume days in April, August, and September 1999, and in January 2000 all occur at the bottom of a price move. Prices reverse direction immediately after the spike, and usually that reversal is substantial. A volume spike does not indicate the strength of the price reversal, it simply tells you that the current move is exhausted. It may turn out to be a major top or bottom, or simply a local turning point.

Volume spikes are a good example of extremes and the clearest cases for trading. The theory of a spike is that, when everyone has entered the market, there is no one left to buy (or sell) and prices must reverse. The crowd is always wrong—at least their timing is always wrong.

Drop in Volume

Although a drop in volume is less impressive than a volume spike, it can be equally important. Volume can decline because there is little interest in a stock or futures market, which often happens when prices are very low. Volume can also drop when a price

FIGURE 12.1   AOL price and volume showing three major volume spikes and declining volume while prices rise at the end of 1999.

reaches *equilibrium,* the price at which buyers and sellers agree is *fair value.* Volume may also drop on the day before a holiday, or just by chance.

While there can be seasonal and other predictable patterns associated with a decline in volume, they all represent a lack of conviction, or a perceived lack of potential price movement, on the part of the traders. In Figure 12.1 there is a steady decline in the average volume from October 1999 through the end of 1999 while prices moved steadily higher. This period had only one-half the volume traded earlier in the year and is interpreted as a lack of investor support for higher prices. Having been hurt earlier in the year, more investors stood aside on the second rally and, while AOL prices did top the previous highs, another sell-off followed. This declining pattern fits the standard interpretation of volume.

STANDARD INTERPRETATION

The interpretation of volume has been part of the trading culture from its beginning. Volume is always considered in combination with price movement:

Volume	Price	Interpretation
Rising	Rising	Volume confirms price rise
Rising	Falling	Volume confirms price drop
Falling	Rising	Volume indicates weak rally
Falling	Falling	Volume indicates weak pullback

This interpretation implies that *volume confirms direction*. When volume declines, it indicates that a change of direction should follow because there is no general support for the price move. Therefore, *volume normally leads price*. It has also been said that 1-week returns accompanied by declining volume tend to reverse the next week. Using weekly data would smooth the volume and may well add consistency.

The classic interpretation of volume was published in the monograph by W. L. Jiler, *Volume and Open Interest: A Key to Commodity Price Forecasting*, a companion piece to his most popular *Forecasting Commodity Prices with Vertical Line Charts*.

In both the futures and stock markets, volume has the same interpretation: When volume increases, it is said to confirm the direction of prices. Price changes that occur on very light volume are less dependable for indicating future direction than those changes associated with relatively heavy volume. An additional uncertainty exists for stocks that are not actively traded, and for low-priced shares where the total dollar volume can be small. In these cases, it might be best to look at the accumulated volume of similar companies or its sector.

Volume and Open Interest

Open interest is a concept unique to futures markets, but helps to explain the depth of the market as well as trader expectations. New *interest* in a market is the result of new buyers and sellers meeting, which increases the open interest, the net of all outstanding contracts being traded. The following table explains the combinations of buyers and sellers that changes open interest:

Buyer	Meets	Seller	Change in Open Interest
New		New	Increase in open interest
New		Old	No change in open interest
Old		New	No change in open interest
Old		Old	Decrease in open interest
Where	"New" buyer (seller) is a trader with no market position seeking to be long (short).		
	"Old" buyer (seller) is a trader who had previously entered a short (long) position and seeks to exit.		

When the open interest increases while prices rise quickly, it is commonly interpreted as more traders entering long positions. This may seem strange because for every new buyer of a futures contract there must be a new seller; however, the seller is likely to be someone looking to hold a position for a few hours or days, looking to profit from the normal ups and downs of price movement. The position trader, who is willing to sit for much more time holding a long position, is the one who is attributed with the open interest. In reality, no one knows. However, if prices keep rising, the shorts are more likely to be forced out while the longs have staying power.

The traditional interpretation of changes in volume and open interest (for futures markets) can be summarized as:

Volume	Open Interest	Interpretation
Rising	Rising	Confirmation of trend
Rising	Falling	Position liquidation (at extremes)
Falling	Rising	Slow accumulation
Failing	Falling	Congestion phase

Three generally accepted notions for the use of volume and open interest are:

1. Open interest increases during a trending period.

2. Volume may decline but open interest builds during an accumulation phase. Volume occasionally spikes.

3. Rising prices and declining volume/open interest indicate a pending change of direction.

There is a traditional interpretation for the combined movement of price direction, volume, and open interest.

Prices	Volume	Open Interest	Interpretation
Rising	Rising	Rising	New buyers are entering the market.
Falling	Falling	Falling	Longs are being forced out; the downtrend will end when all sellers have liquidated their positions.
Rising	Falling	Falling	Short sellers are covering their positions causing a rally. Money is leaving the market.
Falling	Rising	Rising	New short selling. Bearish money is entering the market.

Exceptions

No method is without exceptions, including volume patterns in the stock market. There are days or periods when volume is expected to change, and must be considered in the analysis. For example, volume is expected to decline:

On the first day of the week.
On the day before a holiday.
During the summer.

The most important exception to rising prices and rising volume is the volume spike, which signals a change of direction rather than a confirmation.

Volume is also higher on *triple witching day* (or *quadruple witching day*, if you include futures on stocks), when S&P futures, options on futures, and options on the individual stocks all expire at the same time.

In the futures markets, there are similar patterns. Lighter volume exists during holiday periods and summer months, but may be heavier on Fridays and Mondays during a trending market or a weather market for agricultural products (uncertainty over rain, drought, or frost over the weekend). Liquidation often occurs before the weekend and positions are reentered on the first day of the next week.

Richard Arms' *Equivolume*

Most of the techniques for using volume discussed in this chapter will multiply or accumulate volume, creating an index that rises faster as volume increases. *Equivolume*, a charting method introduced by Richard Arms,[1] takes the unique approach of substituting volume for time along the bottom scale of a chart. When volume increases, the price bar is elongated to the right; therefore, an upwards move on high volume will appear as a higher box that is also wider.

There are no systematic methods for using Equivolume included here and its applications parallel standard chart interpretations; however, analysts can represent this type of chart by creating a new price series in which the daily closing price is repeated based on the relative volume. For example, if the normal volume day causes a closing price to be repeated 10 times, then a day with twice the volume will repeat that price 20 times, and a day with half the volume will show 5 prices. This approach may cause a short-term moving average trend to look strange; however, a linear regression or long-term trend should reflect the importance of varying time.

Herrick Payoff Index

The *Herrick Payoff Index* (HPI) is the most popular calculation that combines price, volume, and open interest. It is designed to gauge the strength of the current trend, and has been applied primarily to futures prices. Because it adds volume and open interest to a momentum calculation, it can be found in Chapter 9 along with some other indicators that use volume as an added feature.

Volume Is a Predictor of Volatility

Most often high volume and high volatility occur at the same time. It is easy to see on a chart that one confirms the other. However, not all days that have high volume also have high volatility. Even on days with high volume, the price can close nearly unchanged from the previous day. We need to look at those days as a sign of *potential volatility*—a large number of traders all with their own objectives somehow managed to offset each other. Tomorrow, if there is an imbalance in the buyers and sellers, and volume is still high, prices could break out in either direction. Therefore, *high volume means high risk*, even on those days when the risk does not materialize.

---

[1] See www.armsinsider.com/education/armsonthemarket/equiv_chart.asp for examples.

VOLUME INDICATORS

Both the stock and futures markets have numerous indicators that use only volume, or those that add volume, intending to make other calculations more robust. The following section gives the most popular of these indicators, most of which originate in the stock market; many use the number of advancing and declining stocks. Readers should note the way in which the data is used from one technique to another and consider the significance of these changes. They are discussed at the end of this section.

Average Volume

The most basic of all volume indicators is the *average*, and the calculation most commonly used for the equity markets is 50 days, although it may be reasonable to use the same period as the price average that is being used. Then, if you are tracking the 200-day moving average, a 200-day average of the volume would make sense. Anything less than 50 days is not likely to be smooth.

Normalizing the Volume

It is always convenient to normalize prices to view them compared to the recent average. This can also be done for volume. Letting the calculation period, *N*, be either 50 or 200 days, we can normalize the volume and represent the result as a percent,

$$NormVol_t = 100 \times \frac{V_t}{\left(\sum_{i=t-P+1}^{t} V_t\right)/N}$$

The normalized volume lets us say that "today's volume is 20% higher than the volume over the past 200 days." This can be seen in the third panel of Figure 12.2, showing Microsoft prices and volume.

Volume Momentum and Percentage Change

Two other basic volume indicators are *momentum* and *rate of change*. These techniques treat volume as they would price. For momentum, this means finding the change in volume over a specific time interval; percentage change measures the size of the volume change relative to the starting value. If *t* is today and *n* is the number of days back (the observation period), then

$$Volume\ momentum_t = Volume_t - Volume_{t-n}$$

$$Volume\%_t = \frac{Volume_t - Volume_{t-n}}{Volume_{t-n}}$$
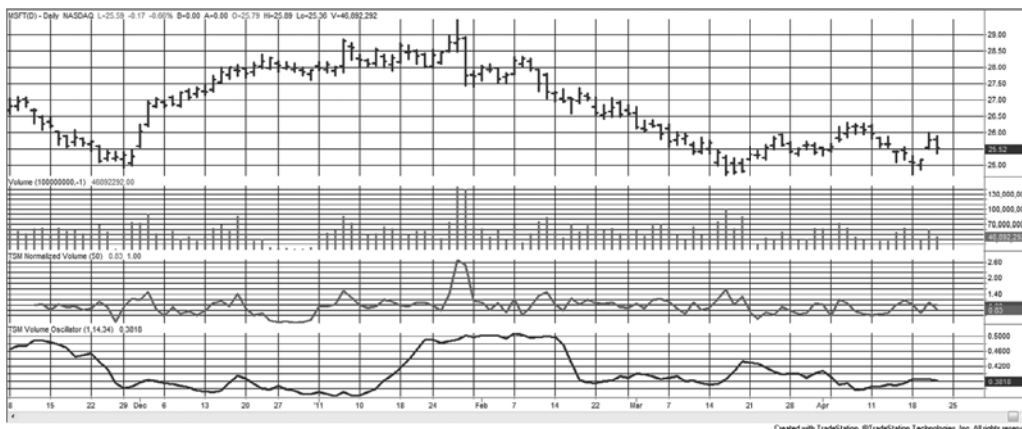
FIGURE 12.2   Microsoft prices (top), volume (second panel), normalized 50-day volume (third panel), and volume oscillator from ratio (bottom), based on 14- and 34-day periods.

Because of the high variance in volume from day to day, volume momentum tends to increase the erratic pattern. It will be necessary to smooth the volume momentum in order to have a useful indicator.

Force Index

Devised by Alex Elder, the *Force Index* is the change in price multiplied by the daily volume,

$$Force\ Index_t = (Close_t - Close_{t-1}) \times Volume_t$$

The Force Index will be positive or negative if the price change was higher or lower. It is also not limited to daily charts, but can be applied to weekly or monthly data, as can many of the other methods in this chapter. Once the daily Force Index is calculated, it should be smoothed. For timing entries and exits, it is suggested that a 2-day exponential moving average (EMA) with a smoothing constant 0.667 be used to avoid unnecessary noise in the indicator. For longer-term analysis, a 13-day EMA (smoothing constant 0.1428) should be used. Consistent with the idea of noise, the 2-day smoothed Force Index *buys* when the value is low and *sells short* when it's high; however, the 13-day value is treated as a trend, giving a buy signal when it crosses above zero and a sell when it crosses below zero. Signals remain in effect until the smoothed Force Index crosses the zero line in the other direction.

Volume Oscillator

Visualizing the pattern of volume can be very helpful, and the simplest way to do this is with a *volume oscillator*, which addresses the issue of erratic data by using two smoothed values. A volume oscillator is a two-step process and may use trends of any calculation period. Here, 14- and 34-day periods are used.

Method 1

Calculate the difference between a short-term and long-term average of volume. Using 14 and 34 days gives

$$VO(Diff)_t = \frac{\sum_{i=t-14+1}^{t} V_i}{14} - \frac{\sum_{i=t-34+1}^{t} V_i}{34}$$

Method 2

Calculate the ratio of a short-term and long-term sum of volume. Again, using 14 and 34 days,

$$VO(Ratio)_t = \frac{\sum_{i=t-14+1}^{t} V_i}{\sum_{i=t-34+1}^{t} V_i}$$

Once calculated, the oscillator values can be displayed as a histogram or a line chart. Figure 12.2 shows Microsoft from November 2010 through April 2011 in the top panel, volume in the second panel, the normalized 50-day volume in the third panel, and the volume oscillator using the ratio at the bottom. In this example, the volume oscillator does not confirm the upwards or downwards trends. Perhaps that is the correct interpretation because these trends were short lived.

On-Balance Volume

Made famous by Joseph Granville, *On-Balance Volume*[2] is now a byword in stock analyst circles. On days when prices close higher, it is assumed that all volume represents the buyers; on lower days, the volume is controlled by the sellers. The volume is then added or subtracted from a cumulative value, *OBV*, to get today's value.

IF *Today's price change* $(p_t > p_{t-1})$ then $OBV_t = OBV_{t-1} + Volume_t$

IF *Today's price change* $(p_t < p_{t-1})$ then $OBV_t = OBV_{t-1} - Volume_t$

The result is a volume-weighted series, which serves to smooth out the erratic nature of the data. In the following formula, the expression in parentheses simply uses the closing prices to produce a value of +1, 0, or −1, which then determines whether today's volume will be added or subtracted from the volume series:

$$OBV_t = OBV_{t-1} + \frac{Close_t - Close_{t-1}}{|Close_t - Close_{t-1}|} \times Volume_t$$

---

[2] Robert W. Colby and Thomas A. Meyers, *The Encyclopedia of Technical Market Indicators* (New York: McGraw-Hill, 2003), is a comprehensive study of most market indicators, including On-Balance Volume and some other techniques in this section.

TABLE 12.2 Calculating On-Balance Volume

| Day	Closing Price	Daily Volume (in 1000s)	On-Balance Volume
1	310	25	0
2	315	30	30
3	318	27	57
4	316	15	42
5	314	12	30
6	320	28	58

Determining the OBV manually is a simple accumulation process as shown in Table 12.2. On day 2 prices closed higher; therefore, the volume of 30 (assume thousands) is added to the starting volume of zero. On the next day, the price was again higher, and the volume is added to the total. On day 4 the price drops so that day's volume is subtracted from the total. In this example, volume was higher when price closed up and lower when prices closed down. This is the ideal pattern in a market with a clear uptrend. Traditionally, the advantage of recording the OBV is in observing when the trend of the prices diverges from the OBV values. The general interpretation of OBV is given in Table 12.3.

Because a volume series has many erratic qualities caused by large variations in volume from day to day, it is most often used with a moving average trend of up to 100 weeks, then identifying a simple volume direction when the OBV value crosses the trend. An upward trend in volume is interpreted as a confirmation of the current price direction, while a downturn in volume can be liquidation or uncertainty. It is intended that the OBV values be used instead of price for making trading decisions.

In Figure 12.3, the OBV is plotted along with the volume at the bottom of the GE chart, covering from December 2010 through April 2011. The pattern of the OBV is similar to the pattern of prices in the top panel even though it is created from volume. The primary difference between the price chart and the OBV line is that the peak in the center is

TABLE 12.3 Interpreting On-Balance Volume

Price Direction	OBV Direction	Interpretation
Up	Up	Clear uptrend
	Sideways	Moderate uptrend
	Down	Weak uptrend near reversal
Sideways	Up	Accumulation period (bottom)
	Sideways	No determination
	Down	Distribution period (top)
Down	Up	Weak downtrend near reversal
	Sideways	Moderate downtrend
	Down	Clear downtrend
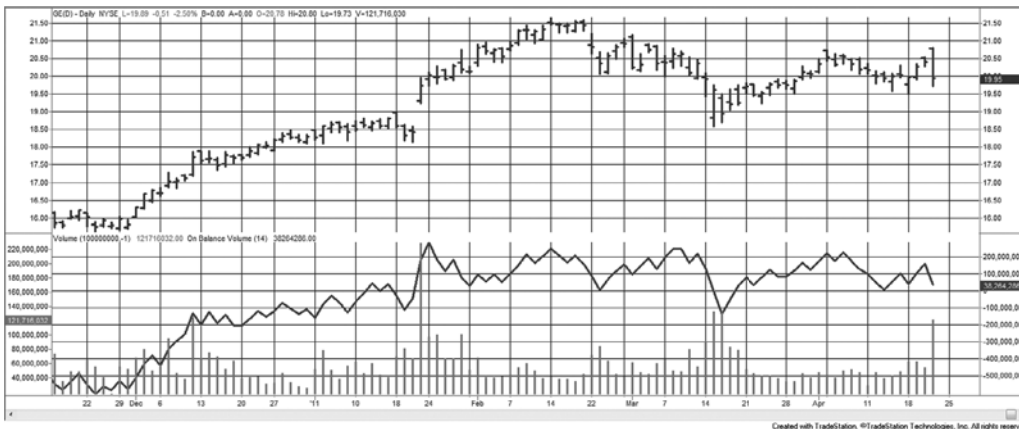
FIGURE 12.3   On-Balance Volume applied to GE from December 2010 through April 2011. Volume is shown on the left lower scale and the OBV on the right lower scale.

flattening and the OBV does not decline as far. This divergence in the center of the chart, prices moving higher than the OBV, would be interpreted as a divergence in which prices will stop rising.

The spreadsheet code, *TSM OBV*, is available on the Companion Website.

Substituting the OBV for Price in a Moving Average

If the OBV systematically improves, then we can use the OBV instead of price to find trend signals. In Figure 12.4, the S&P *e*mini prices are shown in the top panel with a 40-day moving average, and the volume, OBV, and a 40-day moving average of the OBV are in the bottom panel. The corresponding moving average trendlines are equally smooth, but slightly different on the far right where prices rise sooner than the OBV.

FIGURE 12.4   S&P prices are shown with the volume, OBV, and a 40-day trend of the OBV in the bottom panel. Trading signals are derived from the OBV trendline. A corresponding 40-day moving average is shown in the top panel.

It would be necessary to do a comprehensive test on many markets to find out if the OBV produced better trend results than simply using the price. We already know that the S&P and stocks in general are not the best choice for trending; therefore, tests of interest rates would be more interesting. Volume indicators that combine the price move with volume are additional possibilities. An example of a strategy, *TSM OBV Trend*, that applies a trend to a volume indicator can be found on the Companion Website.

Money Flow Index

A simple variation of On-Balance Volume uses the average of the high, low, and closing prices instead of only the close. It is called the *Money Flow Index*, and has been compared to the RSI.

$$TP_t = (H_t + L_t + C_t)/3$$

$$\text{If } TP_t > TP_{t-1} \text{ then } MFP_t = MFP_{t-1} + TP_t \times V_t$$

$$\text{If } TP_t < TP_{t-1} \text{ then } MFN_t = MFN_{t-1} + TP_t \times V_t$$

$$MFI_t = 100 - \frac{100}{\left(\dfrac{1 + MFP_t}{MFN_t} \right)}$$

Volume Count Indicator

An indicator that closely resembles On-Balance Volume is a running total of the days when volume increases minus the days when volume declines. That is, add 1 to the cumulative value on a day when today's volume is greater than the previous day; otherwise, subtract 1. This *Volume Count Indicator* (VCI) can be written as:

$$VCI_t = VCI_{t-1} + \frac{V_t - V_{t-1}}{|V_t - V_{t-1}|}$$

Volume Accumulator

A variation on Granville's OBV system is Mark Chaiken's *Volume Accumulator* (VA). Instead of assigning all the volume to either the buyers or the sellers, the Volume Accumulator uses a proportional amount of volume corresponding to the relationship of the closing price to the intraday mean price. If prices close at the high or low of the day, all volume is given to the buyers or sellers as in the OBV calculation.[3] If the close is at the midrange, no volume is added. This can be accomplished with the following calculation:

$$VA_t = VA_{t-1} + \left(\frac{Close_t - Low_t}{High_t - Low_t} - 0.50 \right) \times 2 \times V_t$$

---

[3] Both OBV and the Volume Accumulator are characterized as "momentum" systems. See Chapter 9 for more information on similar techniques.

Accumulation Distribution

Similar to the Volume Accumulator, and also developed by Mark Chaiken, the *Accumulation Distribution* indicator as well as the Intraday Intensity, the Price and Volume Trend, and the Positive and Negative Volume Indexes use a concept that relates to buying and selling pressure. This compares the strength of the close compared to the open divided by the trading range. It is also called *money flow* in some references. In the following formula, the previous close can be substituted for the open if the opening price is not available.

$$AD_t = AD_{t-1} + \frac{C_t - O_t}{H_t - L_t} \times V_t$$

Intraday Intensity

Intraday Intensity uses the distribution of daily trading range to create a new indicator. If the intraday low is further from the close than the intraday high, in other words, prices close nearer to the high, then the indicator increases in value.

$$II_t = II_{t-1} + \left(\frac{(C_t - L_t) - (H_t - C_t)}{H_t - L_t} \right) \times V_t$$

Price and Volume Trend, the Positive Volume Index, and the Negative Volume Index

The *Price and Volume Trend* (PVT)[4] applies volume to the daily percentage price change, which can be positive or negative. Note that this cannot be used for back-adjusted futures prices, only for cash markets, stock prices, and indexes.

$$VPT_t = VPT_{t-1} + \left(\frac{C_t}{C_{t-1}} - 1 \right) \times V_t$$

There are two variations on the *Positive Volume Index* (PVI) and *Negative Volume Index* (NVI). Both take the approach that a single indicator that adds and subtracts volume based on market direction is not as informative as two separate series that can be viewed at the same time. In the first variation, the direction of the closing price determines whether volume is added to the positive or negative index. In the second variation, higher or lower volume, compared to the previous day, is the basis for adding to the positive or negative index.

---

[4] Dennis D. Peterson, "Positive Volume Index," *Technical Analysis of Stocks & Commodities* (April 2003). Also see Norman G. Fosback, *Stock Market Logic* (The Institute for Economic Research, 1985).

Variation 1, the close as the determining factor:

$$\text{If } C_t > C_{t-1} \text{ then } PVI_t = PVI_{t-1} + V_t$$

$$\text{If } C_t < C_{t-1} \text{ then } NVI_t = NVI_{t-1} + V_t$$

Variation 2, the volume as the determining factor:

$$\text{If } V_t > V_{t-1} \text{ then } PVI_t = PVI_{t-1} + \frac{C_t}{C_{t-1}} \times V_t$$

$$\text{If } V_t < V_{t-1} \text{ then } NVI_t = NVI_{t-1} + \frac{C_t}{C_{t-1}} \times V_t$$

In Norman Fosback's *Stock Market Logic*, the author studied stock trends from 1941 through 1975 based on variation 1 and concluded that:

- If the *PVI* trend is *up* there is a 79% chance that a bull market exists.
- If the *PVI* trend is *down* there is a 67% chance that a bear market exists.
- If the *NVI* trend is *up* there is a 96% chance that a bull market exists.
- If the *NVI* trend is *down* there is a 50% chance that a bear market exists.

In order to find the trend of either the *PVI* or *NVI*, a 6-month (127-day) or 1-year (255-day) moving average was applied to the individual index values.

Aspray's Demand Oscillator

Using direction to separate volume into two series of *Buying Pressure* and *Selling Pressure*, Aspray then nets them into his own *Demand Oscillator*.[5] Note that during a rising market the Selling Pressure has been *divided* by a percentage of the volume, which has been scaled to be greater than 1. The following calculations show the separate steps needed to create Aspray's *Demand Oscillator*:

For rising prices:

$$(Buying\ Pressure)\,BP_t = V_t$$

$$(Selling\ Pressure)\,SP_t = \frac{V_t}{K \times \dfrac{C_t - C_{t-1}}{C_{t-1}}}$$

[5] Thomas Aspray, "Fine-Tuning the Demand Index," *Technical Analysis of Stocks & Commodities* (June 1986), and "Demand Oscillator Momentum," *Technical Analysis of Stocks & Commodities* (September 1989).

For declining prices:

$$BP_t = \frac{V_t}{K \times \dfrac{C_t - C_{t-1}}{C_{t-1}}}$$

$$SP_t = V_t$$

and where

$$K = \frac{3 \times C_t}{\left[\sum_{i=t-9}^{t}\left(\max\left(H_t, H_{t-1}\right) - \min\left(L_t, L_{t-1}\right)\right)\right]/10}$$

and

$$Demand\ Oscillator_t = BP_t - SP_t$$

K is a volatility scaling factor that is 3 times the closing price divided by the 10-day moving average of the maximum 2-day high-low combined range, and has a value likely to be well over 100. For example, if the S&P is trading at 1,000 and the average volatility is 15 full points (which is very high for a 10-day average), then $K = 3000/15 = 200$.

Tick Volume Indicator

In a manner similar to Wilder's RSI, Blau double-smoothes the tick volume as a way of confirming price direction. The *Tick Volume Indicator*, seen in Figure 12.5, is calculated as

$$TVI(r,s) = \frac{100 \times DEMA(upticks, r, s) - DEMA(downsticks, r, s)}{DEMA(upticks, r, s) + DEMA(downsticks, r, s)}$$

TVI ranges from −100 to +100 and *DEMA* (Double Exponential Moving Average) is the double smoothing of the downticks or upticks. The exponential smoothing is first calculated over r bars, and the result of that smoothing is again smoothed over the past s bars. More about this double-smoothing technique can be found in Chapter 7. This technique differs from Blau's price smoothing because it does not first create a momentum series; therefore, the *TVI* will be lagged slightly less than half the sum of the two calculation periods.

Volume-Weighted MACD

The MACD is based on the difference between two moving averages (see Chapter 9). To add volume, the closing prices for each day are multiplied by the corresponding volume, and each moving average is normalized by the average volume over the same period. The recommended calculation periods are 12 and 26 days, the standard MACD values. A
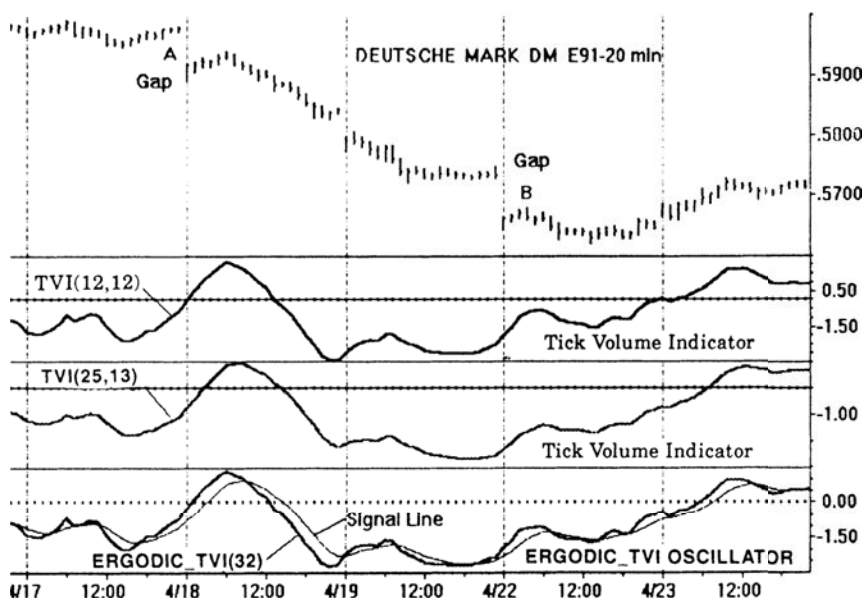
FIGURE 12.5 Blau's Tick Volume Indicator.

Source: William Blau, *Momentum, Direction, and Divergence* (New York: John Wiley & Sons, 1995), 45.

signal line, also the same as MACD, is the exponentially smoothed VWMACD line using a 0.20 smoothing constant, the equivalent of about 9 days.

$$VWMACD_t = \left(\frac{\sum_{t=t-12+1}^{t} C_i \times V_i}{\sum_{i=t-12+1}^{t} V_i} \right) - \left(\frac{\sum_{t=t-26+1}^{t} C_i \times V_i}{\sum_{i=t-26+1}^{t} V_i} \right)$$

Figure 12.6 uses a program *TSM Volume Weighted MACD*, available on the Companion Website, to plot the VWMACD for euro currency futures, which appears in the bottom panel. The heavier line is the VWMACD, and the thinner, smoother line is the signal. As with the MACD, the rules are to enter a new long position when the VWMACD crosses above the signal line, and the opposite for shorts. When compared to the MACD, this variation appears to make only small changes, yet a slight improvement in timing can be a great advantage in trading.

Variably Weighted Moving Average Using Volume

A unique approach taken by Christian Fries[6] uses the relationship between the number of outstanding shares, the current price, and the volume of the next period or the next

---

[6] Christian P. Fries, "Elastic Moving Averages," *Technical Analysis of Stocks & Commodities* (June 2001).
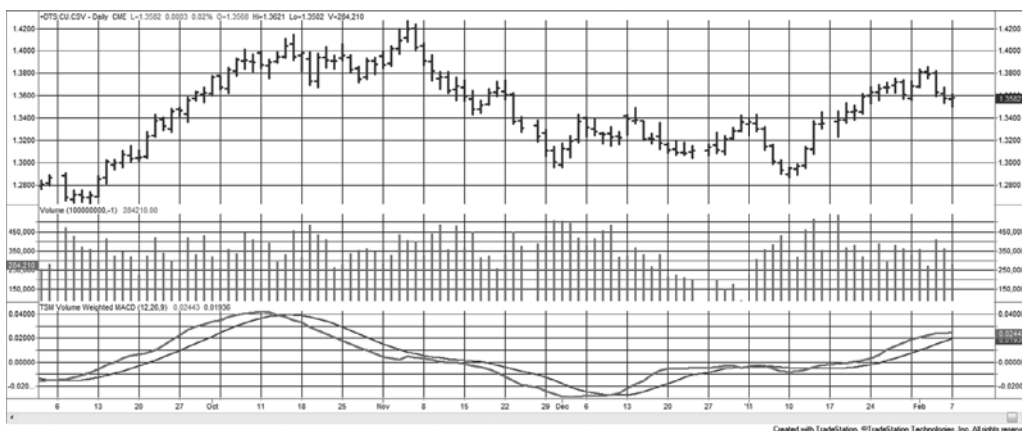
FIGURE 12.6 Volume-weighted MACD applied to euro currency futures.

trade, to create an *Elastic Volume-Weighted Moving Average* (*eVWMA*). Using the difference between the number of outstanding shares and the number of shares being traded he creates a weighting factor. That weighting factor causes the previous trend value to have more importance when fewer shares are traded and less weight when relatively more shares are traded. The net effect is that the weighted average is more responsive to change when relatively more shares are traded.

$$eVWMA_t = \frac{(OS - V_t) \times eVWMA_{t-1} + V_t \times C_t}{OS}$$

where OS = the number of outstanding shares.

Using Merck (MRK) as an example, the total outstanding shares was set to 200 million, based on an observed maximum daily volume of approximately 150 million. By keeping the outstanding shares slightly bigger than the maximum volume, the indicator will vary across a wider range and show more of its characteristics. Figure 12.7 compares the eVWMA to a standard exponential smoothing of 0.10. The results are very similar although the eVWMA tracks closer to the price in the center of the chart.

Substituting Open Interest for Volume Using Futures

Futures volume can create very different results from volume for stocks. The markets themselves—energy, gold, and even interest rates—have been in an uptrend for years. This causes many of the indicators, especially OBV, to move steadily higher, while it appears to oscillate when applied to stocks over the past 10 years.

It has already been mentioned that futures volume comes in two parts, contract volume and total volume. Contract volume begins at zero, increases as the contract moves from deferred (many months away from delivery) to the nearby delivery, then back to zero as the contract expires. Open interest follows a similar pattern. To avoid this pattern, total
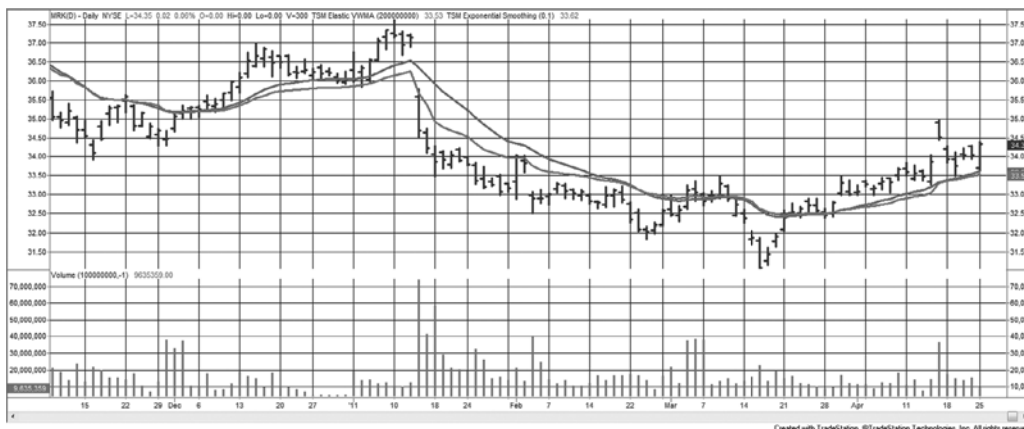
FIGURE 12.7 Elastic Volume-Weighted Moving Average applied to Merck (MRK) compared with a 0.10 exponential smoothing.

volume is used instead of contract volume. In the same way, futures data shows contract open interest and total open interest. That data varies in the same pattern as volume; however, open interest is much more stable than volume. Similar to volume, open interest increases as market activity increases; therefore, it reflects the same fundamentals.

Total open interest can be a preferable substitute for volume in many of the indicators covered in the previous sections. It should be remembered that open interest, as well as volume, is usually available on downloaded data one day late. These numbers could be manually downloaded from Bloomberg, but it is safer to use them as a longer-term indicator, which best suits their purpose.

A spreadsheet example of applying OBV to S&P futures, using both total volume and total open interest, is available on the Companion Website as *TSM OBV*.

VWAP

As a note, order executions can be placed as a *VWAP*, a *volume-weighted average price*. For larger positions, this would return an average price representing how actively the market traded at different levels throughout the day. It is most convenient for hedge funds that do not want to force prices higher by placing an excessively large buy order at one point during the day, or even on the close. Equally spaced orders that ignore liquidity use *TWAP*, a *time-weighted average price*.

BREADTH INDICATORS

Market breadth measures the imbalance between the number of advancing and declining stocks on a given day. It is the percentage of rising stocks to the total number of stocks traded. In general, more advancing issues should add confidence to an upwards price

move in the same way that volume confirms price. A net increase in the S&P 500 while more issues are declining should generate concern that the upwards move is poorly supported.

The information for advances and declines is published along with stock prices each day. You may also want the volume of advancing issues and the volume of declining issues. Both of these values can be found in the *Wall Street Journal* under "Trading Activity." In this section, the CSI data series, UVDV, will be used. In addition, we will look at the new highs and new lows using CSI data MKST, to see if those numbers add value to price analysis.

To interpret the CSI data files, use the following guide. The normal field designation is in the heading and the actual content is below.

Series	Date	Open	High	Low	Close	Volume	OpInt
UVDV	Date	Total traded	Advanced	Declined	Unchanged	Up Volume	Down Volume
MKST	Date	Total traded	Advanced	Declined	Unchanged	52-Week Highs	52-Week Lows

Market breadth indicators are typically used as a confirmation of price direction. The table below shows that the interpretation of market breadth is the same as the relationship between volume and price.

Market Breadth	Price	Interpretation
Rising	Rising	Breadth confirms price rise
Falling	Falling	Breadth confirms price drop
Falling	Rising	Breadth does not confirm price rise
Rising	Falling	Breadth does not confirm price drop

The measurement of breadth must relate to what you are trading. If 75% of the stocks in the Dow rise, the S&P Small Cap Index can still be lackluster or declining. Investors shift from small-cap stocks to the S&P 500 and then to the Dow when they are looking for safety. The breadth of the overall market, or the breadth of large caps, is not enough to confirm an upwards trend in the small caps.

It is worth noting that Gerald Appel, a well-known market analyst and technician, points out that, when using market breadth, it is best if it is equally weighted (i.e., a simple count).

Separating Advancing from Declining Volume

The most basic of all techniques for determining market breadth is the *Advance-Decline Index*, which creates a new series by adding the net of the number of advancing and declining stocks each day. Unchanged issues are ignored.

$$Advance\text{-}Decline\ Oscillator_t = Advances_t - Declines_t$$

$$Advance\text{-}Decline\ Index_t = Advance\text{-}Decline\ Index_{t-1} - Advances_t - Declines_t$$

In some of the techniques that follow, only the oscillator, or single-period calculation, may be shown. An index can be created by accumulating these values.

Sibbett's Demand Index

One method that smoothes volume by using the total activity of the past 10 days is James Sibbett's *Demand Index*. The technique bears a resemblance to the approach used in Wilder's RSI. It can be used as an oscillator with individual daily values or accumulated into an index. Where $n = 10$, the individual days are calculated as the sum of the upside volume divided by the sum of the downside volume:

$$Demand\ Index_t = \frac{\sum_{i=t-n+1}^{t} upside\ volume_i}{\sum_{i=t-n+1}^{t} downside\ volume_t}$$

The values for upside and downside volume can also be found in the *Wall Street Journal* under the heading "Trading Activity." Arthur Merrill, another well-known stock analyst, has suggested using $n = 5$.

Bolton-Tremblay

Introducing the net advancing stocks compared to the number of stocks that were un-changed, the Bolton-Tremblay approach also includes a form of geometric weighting of the results. In the following calculation, *BT* is the daily Bolton-Tremblay value, and *BTI* is the index resulting from accumulating the individual daily values.

$$BT_t = \frac{Advancing_t - Declining_t}{Unchanged_t}$$

$$\text{if } BT_t > 0 \text{ then } BTI_t = BTI_{t-1} + \sqrt{BT_t}$$

$$\text{If } BT_t < 0 \text{ then } BTI_t = BTI_{t-1} - \sqrt{BT_t}$$

The Bolton-Tremblay calculation will have large spikes on days when the advancing or declining breadth is very strong, which make the unchanged number of stocks in the denominator very small.

Schultz

Schultz chose to look at the advancing stocks only as a percentage of the total stocks, which puts the results in the convenient range of 0 to 100:

$$Schultz\ A/T_t = \frac{Advancing_t}{Total\ stocks_t}$$

McClellan Oscillator

The *McClellan Oscillator* starts with the net advances (NA) and then creates an oscillator by subtracting two smoothed trends, 19 and 39 days, based on the net advances, in a manner similar to the MACD. Converting the 19- and 39-day calculation periods to smoothing constants using $2/(n+1)$ we (conveniently) get $s_1 = 0.10$ and $s_2 = 0.05$.

$$NA_t = Advances_t - Declines_t$$

$$E1_t = E1_{t-1} + s_1 \times \left(NA_t - E1_{t-1}\right)$$

$$E2_t = E2_{t-1} + s_2 \times \left(NA_t - E2_{t-1}\right)$$

$$McClellan\ Oscillator_t = E1_t - E2_t$$

Figure 12.8 compares Sibbett's Demand index with the more complicated McClellan Oscillator. S&P futures prices are in the top panel, the Demand Index in the second panel, the up and down volume in the third panel, and the McClellan Oscillator at the bottom. The scales of the two indicators are very different. The Demand Index is a ratio; therefore, it varies above and below 1.0, and reflects the upside/downside volume over a rolling 10-day period. The McClellan Oscillator uses a longer calculation period and smoothes even further using exponential smoothing.

Both indicators stay slightly above neutral during the long upward move, but the most interesting feature would be the extreme highs and lows, which are clearer in the McClellan Oscillator, corresponding nicely to two price lows and one peak. The high point that ended the long upwards move shows as divergence in the two indicators, a pattern that is not as easy to identify at the time, but can be very useful.
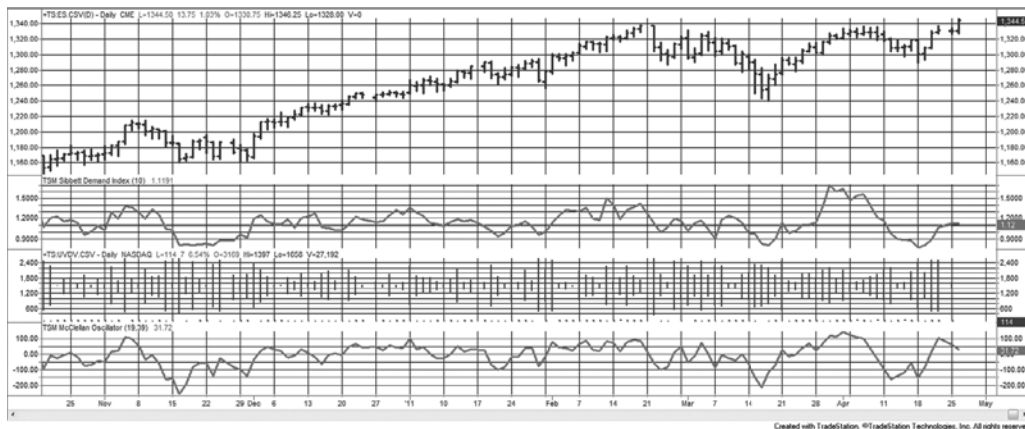
FIGURE 12.8 S&P futures (top panel), Sibbett's Demand Index (second panel), the up/down volume (third panel), and the McClellan Oscillator (bottom panel).

Upside/Downside Ratio

A simple *Upside/Downside Ratio* (UDR) is similar to Sibbett's Demand Index but is not smoothed. With this ratio, it is easier to see the asymmetry of the calculation. When the declining volume is very low, for example only 20% of the total volume, the UDR will have a value of 4.0. In contrast, if the advancing volume is low (20% of the declining volume), the ratio is 0.20. There is no upside limit when stocks are advancing but the downside ratio is limited to the range between 0 and 1. This is a familiar pattern for advancing and declining prices expressed as a percentage. High values of the UDR are expected to precede a bull market period.

$$UDR_t = \frac{Advancing\ Volume_t}{Declining\ Volume_t}$$

Arms Index (TRIN)

Richard Arms explains the relationship between the number of advancing and declining stocks, and the up and down volume, in the *Arms Index*, more popularly known as *TRIN*, the *Trader's Index*. It divides the ratio of the number of advancing and declining stocks by the ratio of the volume of the advancing and declining stocks.

$$TRIN_t = \frac{Advancing\ stocks_t\ /\ Declining\ stocks_t}{Advancing\ volume_t\ /\ Declining\ volume_t}$$

Thrust Oscillator

Tushar Chande's *Thrust Oscillator* (TO) uses the same values as *TRIN* but creates a bullish oscillator by multiplying the number of advancing stocks by the volume of the advancing stocks and subtracting the comparable declining values, then dividing by the sum of the two.

$$TO_t = 100 \times \frac{Adv\ stocks_t \times Vol\ of\ adv\ stocks_t - Decl\ stocks_t \times Vol\ of\ decl\ stocks_t}{Adv\ stocks_t \times Vol\ of\ adv\ stocks_t + Decl\ stocks_t \times Vol\ of\ decl\ stocks_t}$$

The Thrust Oscillator is used as an overbought/oversold indicator to complement another strategy that would decide on the major direction of the market. Values of ±30 may be used to identify overbought and oversold levels, respectively. Whenever an indicator is not smoothed, the extremes become the most interest points of action.

Figure 12.9 shows the Thrust Oscillator (second panel) along with S&P futures (top panel). The bands of ±30 give short-term indications of overbought and oversold days and occur more frequently during a sideways price move. For the 3-month period when the S&P was steadily rising, the indicator gave many more buy signals (penetration of the bottom band) than sell signals, a good outcome.
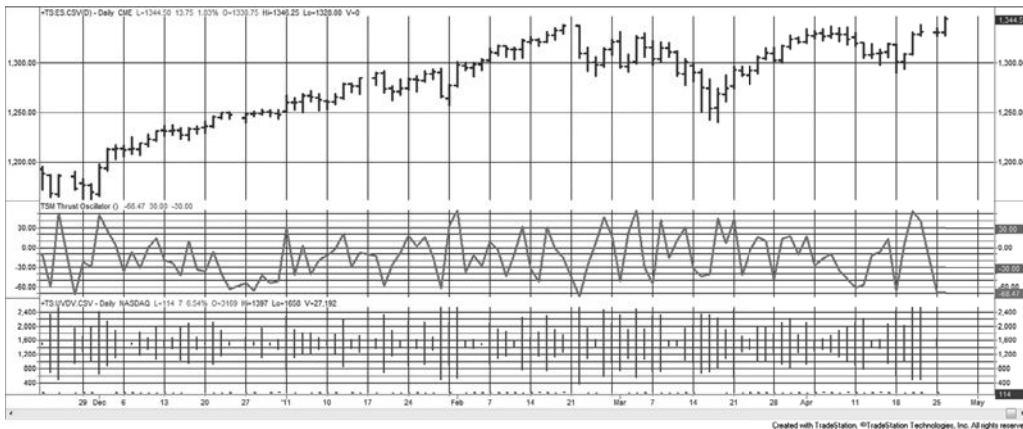
FIGURE 12.9 The Thrust Oscillator (second panel) shown with S&P futures (top panel).

When studying both the Arms Index (TRIN) and the Thrust Oscillator (TO) it would seem sensible to express all the component values as percentages of the whole. That is, the number of advancing and declining stocks should each be divided by the total number of stocks, and the volume of the advancing and declining stocks should each be divided by the total volume of all shares traded. This would make those days with low relative volume less important. In the current calculations, inactive days might produce highly distorted values.

New Highs and Lows

There is very little literature on the use of the number of new highs and lows that occur daily, yet the data are available and can be used in ways similar to market breadth. The most obvious way to represent this data is by creating a *High-Low Index* (HLX),

$$HLX_t = HLX_{t-1} + NH_t - NL_t$$

where *NH* are new 52-week highs and *NL* are new 52-week lows. Another approach would be a simple ratio of the new highs compared to the combined new highs and lows, over the past *n* days,

$$HLR_t = \frac{\sum_{i=t-n+1}^{t} NH_i}{\sum_{i=t-n+1}^{t} (NH_i + NL_i)}$$

Gerry Appel[7] uses this *High-Low Ratio*, smoothed over 10 days, and generates a buy signal when the ratio crosses above a threshold of, for example, 0.80 or 0.90, and a sell

---

[7] Gerald Appel, *Power Tools for Active Traders* (Upper Saddle River, NJ: Financial Times Prentice Hall, 2005), 126.
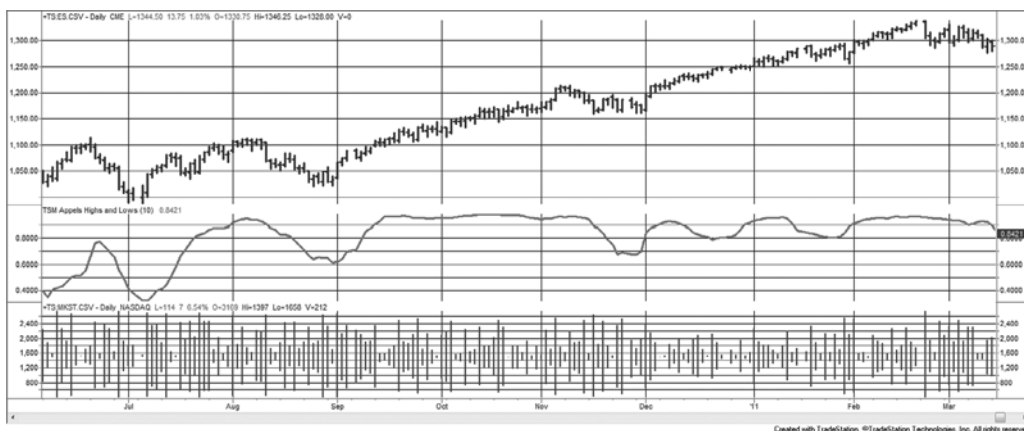
FIGURE 12.10 S&P futures (top panel) shown with the High-Low Ratio (HLR) in the second panel, and the high-low data at the bottom.

signal when it crosses back below that level. In Figure 12.10, the S&P futures are shown with the HLR in the middle panel, and the high-low data in the bottom panel. The top line of the middle panel is 0.90 and might serve as a threshold to extract high-momentum trades during the steady uptrend. It may also be useful to create a fast smoothing of the HLR to use as a signal line, in the same manner as MACD, which would avoid the need for a strict threshold entry and exit value.

Baskets of New Highs and Lows

Another method that will be discussed in more detail in Chapter 13 is buying a basket of stocks making new highs and selling a basket of stocks making new lows. Naturally, this assumes that each stock making a new high will sustain that direction because there has been some fundamental change in its earnings or structure that the market sees as positive. Similarly, those stocks making new 52-week lows have negative fundamental changes. To avoid unnecessary risk, the long basket and short basket must have the same volatility.

Is One Volume or Breadth Indicator Better than Another?

The final test of a successful indicator is whether it adds value to your trading decisions. This can be determined by programming the indicator into a strategy and testing the rules over a reasonable historic time period. Before doing that, we should decide what an increase or decrease in volume or breadth is telling us, and how we will use that information for trading. Volume and breadth indicators are more difficult than those that use price because the values have greater variance. That makes translating a sound concept into a profitable strategy more difficult. If performance testing does not confirm our

expectations, then the indicator needs to be applied differently. There seems no doubt that volume and breadth data are telling us something, and uncovering those messages would improve our trading performance.

Looking at the previous selection of indicators, we see that small net changes in price can result in all volume being designated to one market direction. This is the case with Granville's On-Balance Volume and in the Bolton-Tremblay method. Is it reasonable to add all volume to the accumulated index if the S&P gained only a fraction of one point? On a single day this may seem to be arbitrary, but just as in many basic systems, over a long period of time the net effect of this decision is sound.

The changes made by Mark Chaiken, which take a percentage of volume based on the relative close of prices within the daily range, seem very sensible and avoid the all-or-nothing technique used in On-Balance Volume. Sibbett's Demand Index, by using the sum of 10 days' volume, avoids this problem completely and smoothes out results, a situation often needed when using the highly variable values of volume and breadth.

By seeing these indicators close together, they appear to be a collection of minor manipulation of data. Traditional technicians still advocate interpretation using trend lines, divergence, or new highs and lows. Instead of having to interpret only the difficult movements of prices, they have added an equally difficult series of volume or breadth. While it is clear that volume and breadth contribute to the understanding of a price move, you must first be able to test whether you can use that information profitably. To do that, some fully systematic method is needed to insure that any approach can be implemented and validated.

Breadth as a Countertrend Indicator

All the literature favors using breadth to confirm direction, but Connors argues the opposite.[8] He trades the advance/decline (*AD*) ratio as follows:

Sell the index when *AD* > 2.0
Buy the index when *AD* < 0.50
Exit the trade after 5 days unless a reverse signal occurs

Using the series UVDV-I or MKST-I defined earlier in this section, we added the SPX in column 2, then created the advance/decline ratio. The buy and sell rules were applied, but because there were very frequent signals, all trades were held until an opposite signal occurred, rather than for 5 days. The total PL (in percent) is shown in Figure 12.11. Complete calculations can be found in the Excel spreadsheet, *TSM MKST-I Market Breadth*, on the Companion Website.

---

[8] Larry Connors, "Fade the Breadth," *Futures* (January 2005).
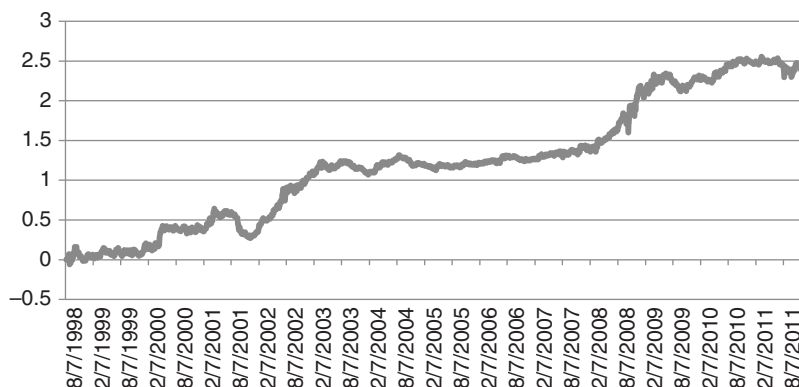
FIGURE 12.11   Total PL using the AD ratio as a countertrend signal.

## INTERPRETING VOLUME AND BREADTH SYSTEMATICALLY

Most systematic approaches to volume apply a long-term smoothing method to the data and then identify trend changes to confirm price direction. This can be implemented with any of the accumulation indexes, but not with single-day momentum or oscillator values. For the oscillators, most analysts have taken the approach that high volume confirms a new price direction; extreme volume, on the other hand, is more likely to be a reversal signal. If volume and price peaks do not occur at the same time, a volume peak should precede a trend change. A decline in volume has also been used to confirm direction, but it is more likely to indicate that prices have reached equilibrium, and that a further advance or decline requires additional confirmation.

Identifying a Volume Spike

A *volume spike* is one of the clearest and best examples of useful volume information. In its most obvious form, a volume spike is a day when the volume is at least twice that of other days over a period generally longer than one month, most likely longer than three months. Volume spikes may come in clusters due to an unstable political or economic situation; however, it is normally associated with the end of a significant price move.

Figure 12.12 shows 12 years of Microsoft prices with volume along the bottom. There are a few dozen times when volume spikes. Just after 2000 there is a cluster of spikes that are associated with the end of the tech bubble, a price drop of 50% in just a few weeks. Most of the spikes on the chart fall on days with large price moves and precede reversals. While the reversals are not always large, the volatility spike marks the end of the current move.

A simple spike can be measured as exceeding a threshold, T, where T is a multiple of the average volume, $AV = average\ (volume, n)$, over the recent n days. Then a volume
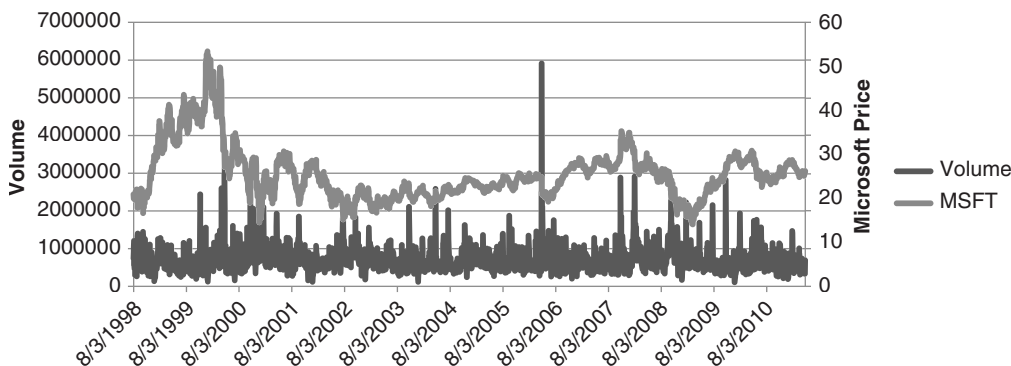
FIGURE 12.12   Microsoft prices with volume along the bottom.

spike exists if today's volume is greater than the average volume on the previous day times its multiplier, $V_t > T \times AV_{t-1}$.

Most cases are not as clear as a 1-day spike following much lower volume. More likely, volume has increased significantly over the past two or three days, culminating in a spike that is much higher than the average volume, but may only be 25% higher than the previous day. The solution to this is to lag the average volume so that it does not reflect the increasing volume of the recent few days.

We will make the assumption that volume resulting in a spike takes no longer than three days to develop. We can safely say that the average volume, lagged five days, should not reflect the recent rise in volume. Then a volume spike would be identified when

$$V_t > \frac{\sum\limits_{i=t-n-5}^{t-5} V_i}{n}$$

The right part of this formula is the average volume over n days beginning $t - n - 5$ days ago and ending $t - 5$ days ago. If t is today, then the average volume stops five days ago; therefore, it is not influenced by recent volume data.

Volume spikes are not as extreme when dealing with futures markets and indexes. In Figure 12.13 NASDAQ futures are shown with total (all contracts) volume for a period in May 2010. Higher volume days correspond to two cases where prices reversed from lows, and the peak in the center of the chart is associated with low volume. Both situations offer an opportunity to improve trading.

In the crude oil chart, Figure 12.14, the spikes in volume can be seen as frequently occurring at price reversal points, while lower volume may be indicative of the continued price direction. Although there are no comprehensive tests available to confirm these observations, high volume seems to precede a price reversal, even if it lasts only one day.
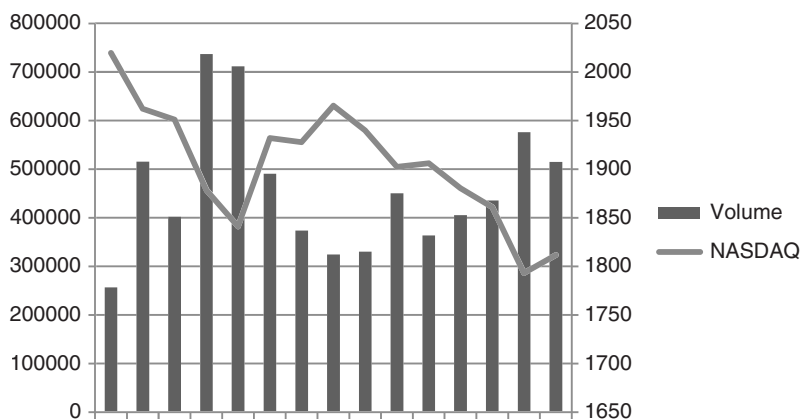
FIGURE 12.13 NASDAQ prices and volume show both high and low volume associated with price reversals.

Moving Average Approaches

Not all analysts agree that higher volume is a reversal signal. A straightforward way of using higher volume to confirm a trend is to calculate a 10-day moving average of the volume as a complement to a 20-day trend position.[9] By simply requiring the current volume to be greater than the average volume over those past 10 days, you introduce the idea of greater participation associated with the new trend. An additional important benefit is that this volume condition acts as a filter, eliminating a substantial number of trades. If the net returns are the same, the volume-filtered approach is far better because you are out of the market more, reducing your risk and not reversing your position every time there is a new signal. That improves liquidity.
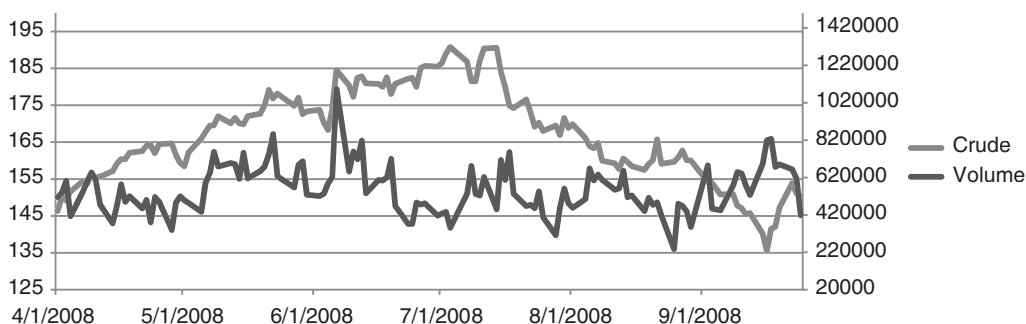
FIGURE 12.14 Crude oil futures (back-adjusted), with total volume shown as a line.

---

[9] Alex Saitta, "A Price and Volume-Based System," *Technical Analysis of Stocks & Commodities* (March 1996).

A similar method was proposed by Waxenberg.[10] A 10-day moving average of the volume is calculated as the normal level, and a change in trend must be confirmed by a 20% increase in volume above this norm. (The 20% threshold acts as an additional smoothing filter, but may be replaced by a longer trend and smaller threshold.) Extremes in a price move can be found at points that exceed approximately a 40% volume increase. Applied to the stock market, Waxenberg used the extreme volumes to indicate the end of a sell-off. To add more flexibility over longer test periods, and to adapt more quickly to volatility changes, Bollinger bands (based on 2 standard deviations, or 95% probability) can be substituted for the fixed percentage thresholds.

Alternately, using 13 days of volume, subtract the total down volume from the total up volume. A plot of the results will serve as a momentum indicator from which over-bought and oversold levels can be identified. If these values are unstable due to lack of liquidity, they may be smoothed using a short-term moving average.

Advance-Decline System

Advance and decline values, as with most volume figures, can be more useful if they are smoothed. By combining peak values of the net of smoothed advancing and declining shares with a directional move in price, Connors and Hayward have created a basic system structure that they named CHADTP[11] (*Connors-Hayward Advance-Decline Trading Patterns*). This system tries to identify reversal patterns by applying the following steps:

1. Add the past 5 days of advancing issues on the New York Stock Exchange.
2. Add the past 5 days of declining issues on the New York Stock Exchange.
3. Subtract the 5-day sum of declining issues (Step 2) from the advancing issues (Step 1).
4. Divide by 5 to get the average daily value.

$$CHADTP = (Sum(AdvancingNYSE,5) - Sum(DecliningNYSE,5)) / 5$$

To trade using this oscillator, Connors and Hayward have determined that ±400 are the extreme levels where the values have been overbought and oversold. Based on this, we can apply the following rules to the S&P futures:

1. *Sell short* when CHADTP > + 400 and the S&P futures trade 10 basis points below the low of the previous day; *buy* when CHADTP < − 400 and the S&P futures trade 10 basis points above the high of the previous day.

---

[10] Howard K. Waxenberg, "Technical Analysis of Volume," *Technical Analysis of Stocks & Commodities* (March 1986).

[11] Laurence A. Connors and Blake E. Hayward, *Investment Secrets of a Hedge Fund Manager* (Chicago: Probus, 1995).

2. Note that the oscillator does not have to exceed its recent extremes on the day of the buy or sell signal.

3. Timing is best if the signal occurs at the same time as a newspaper commentary indicating "depressed volume," or volume significantly below the 3-month average, which is seen as an excess of cash waiting to enter the market.

This system targets returns over a 5- to 7-day period. A drop in the oscillator, which results in values in the midrange, is an opportunity to exit. A standard price oscillator can be constructed to generate overbought and oversold signals within this time frame. An opposite entry signal would reverse the position.

We should be aware that this method was developed in 1995 when the S&P was lower and less volatile. As of 2011, the DJIA is about four times its 1995 value, while the S&P is only about one-third higher. However, the markets are more volatile now. A change of 10 basis points (0.10%) in the S&P is about 1.30 points, which is likely to happen every day. Raising the threshold to 50 or 100 basis points may make it easier to assess the value of this method.

Breadth as a Counter-Trend Indicator

Equities have a tendency to be mean reverting because they have more noise than most other sectors. That seems to hold true for market breadth. Larry Connors has shown that more declining stocks is an excellent leading indicator of a stock market advance.[12]

- If declining NYSE stocks are greater than advancing NYSE stocks for at least three days in a row, then the S&P averaged a gain of 0.50% the following week.
- If declining NASDAQ stocks are greater than advancing NASDAQ stocks for at least two days in a row, then the NASDAQ 100 index gained an average of 0.80% the next week.
- If declining NYSE stocks were at least twice the NYSE advancing stocks, then the S&P gained an average of 0.60% the following week.

In addition, when the opposite occurred, all three situations showed essentially no gain in the respective index during the following week. The observation period was 1996–2003, which contained extreme bull and bear markets.

AN INTEGRATED PROBABILITY MODEL

If there is a noticeable relationship between price, volume, and open interest (or market breadth), then a probability model can be constructed to test its importance.[13] To do this

---

[12] Larry Connors, "Fade the Breadth," *Futures* (January 2005).
[13] Based on the May 1995 "CSI Technical Journal," Commodity Systems Inc., Boca Raton, Florida.

it is necessary to construct a 1-day-ahead forecast using a simple linear regression model and then backtest weighting factors for each element. Because of powerful software products, this has become a very manageable process. Using TradeStation's *EasyLanguage*, each 1-day-ahead forecast is determined using the previous *n* days:

Price forecast: $\qquad Pf = Price + LinearRegSlope(close,\ n)$

Volume forecast: $\qquad Vf = Volume + LinearRegSlope(volume,\ n)$

Open interest forecast: $\qquad Of = Opint + LinearRegSlope(opint,\ n)$

On-Balance Volume forecast: $\qquad OBVf = OBV + LinearRegSlope(OBV,\ n)$

The function *LinearRegSlope* returns the 1-period increase or decrease in the input value based on a straight-line fit of the past *n* days. That value can be added to the current price, volume, open interest, or On-Balance Volume to get the 1-day-ahead forecast. The calculation period *n* can be selected by subtracting the actual next-day value from the forecast and creating an error series that can be measured using a standard deviation. The number of days, *n*, that generates the smallest standard deviation is the best forecast period. It is possible that the optimal forecast period will differ for each of the four items above but a robust solution would seem to be the one period that minimized the errors across all the components. These and other testing issues are thoroughly discussed in Chapter 21.

Having found the four 1-day-ahead forecasts, an index can be created that gives one weight w_1 to the price forecast and the remaining weights, $1 - w_1$, to a combination of the other three factors. This assumes that price is the most important predictor of price. Then

$$Forecast\ index = w_1 Pf + (1 - w_1) \times (w_2 Vf + w_3 Of + w_4 OBVf)$$

This formula can be backtested for values of w_1, w_2, w_3, and w_4 where $\sum w = 1$. The final index can be used instead of price for determining the trend. It will still require a moving average, or some smoothed line, to signal new uptrends and downtrends; however, the results, if successful, should be more reliable than using only price.

One advantage of testing the weighting factors is that, if one of the four elements is not helpful in predicting a trend, the weighting factor should be near zero. Another approach, that does not cluster the nonprice data together, would be to treat each item separately in the classic form:

$$Forecast\ index = w_1 Pf + w_2 Vf + w_3 Of + w_4 OBVf$$

INTRADAY VOLUME PATTERNS

Identifying increases and decreases in volume during the trading day must consider the patterns caused by the flow of orders during the day and the way traders enter orders. During the past five years, there has been a great improvement in timely reporting of

intraday volume for both stocks and futures markets. Services such as Bloomberg have continually updated volume for nearly all markets. Prior to this, only tick volume was the practical alternative. Tick volume represents the number of price changes that occurred during any time interval. It turns out that the frequency of price changes is directly related to the actual volume traded. If you purchase intraday data from CQG, it will have tick volume in each price bar.

Time Stamps

One point to note is that CQG time stamps each bar at the beginning of the interval. It is because they accumulate the data as it occurs, so that 5-minute bars that are stamped 10:00, 10:05, and 10:10 have the data from 10:00:00 through 10:04:59, 10:05 through 10:09:59, and 10:10:00 through 10:14:59, respectively. For many applications, users may want to shift those times to show the end of the bar rather than the beginning. That way, if you are matching the time of the bar with the closing time of the session, those times will match. Otherwise, if the closing time is 4 P.M., or 16:00, the last 5-minute bar will be stamped 15:55. Also note that because these bars are time stamped at the beginning, the last bar of the day will have only the trading that occurred in the last 1 minute of the session. For example, the U.S. bond chart has the last interval posted at 14:00, the Chicago closing time. All trading starting at 14:00 will be captured, which is only the last minute, yet it shows very high volume. You may want to combine the last bar with the previous bar.

Intraday Patterns

Figure 12.15 shows the intraday tick volume patterns for four futures markets, crude oil, the German DAX, NASDAQ, and U.S. 30-year bonds. The charts represent the average number of price changes in each 15-minute interval for the years 1995 through 2005, for the nearest contract to delivery. Each market has a somewhat different pattern.

Before Europe became a viable, accessible market, intraday volume patterns tended to have a very symmetric *U*-shaped pattern, with the highest volume at the open and close. The bottom of the *U* always came at midday when traders took their lunch break. This can still be seen in the NASDAQ chart, Figure 12.15c. As European markets increased in volume the patterns changed so that volume was greatest while both the United States and Europe were open, then declined as European traders ended their business day. In Figure 12.15b, the DAX shows a sharp drop from 17:30, then one strong bar on the close at 18:00 (6 pm). Soon after 2005, the Eurex extended hours to match the U.S. trading sessions, so the DAX now trades until 20:00, the close of the U.S. equity index markets. We should expect the *U*-shaped pattern in the late session to be extended to 20:00.

Crude oil shows two *U*-shaped patterns, the first for the primary session, originally called the *pit session*, and the second for the evening session. Now crude trades 24 hours, although we should always expect this same volume pattern bounded by normal daily business hours in the country where the product is traded. Then the CME Nikkei contract
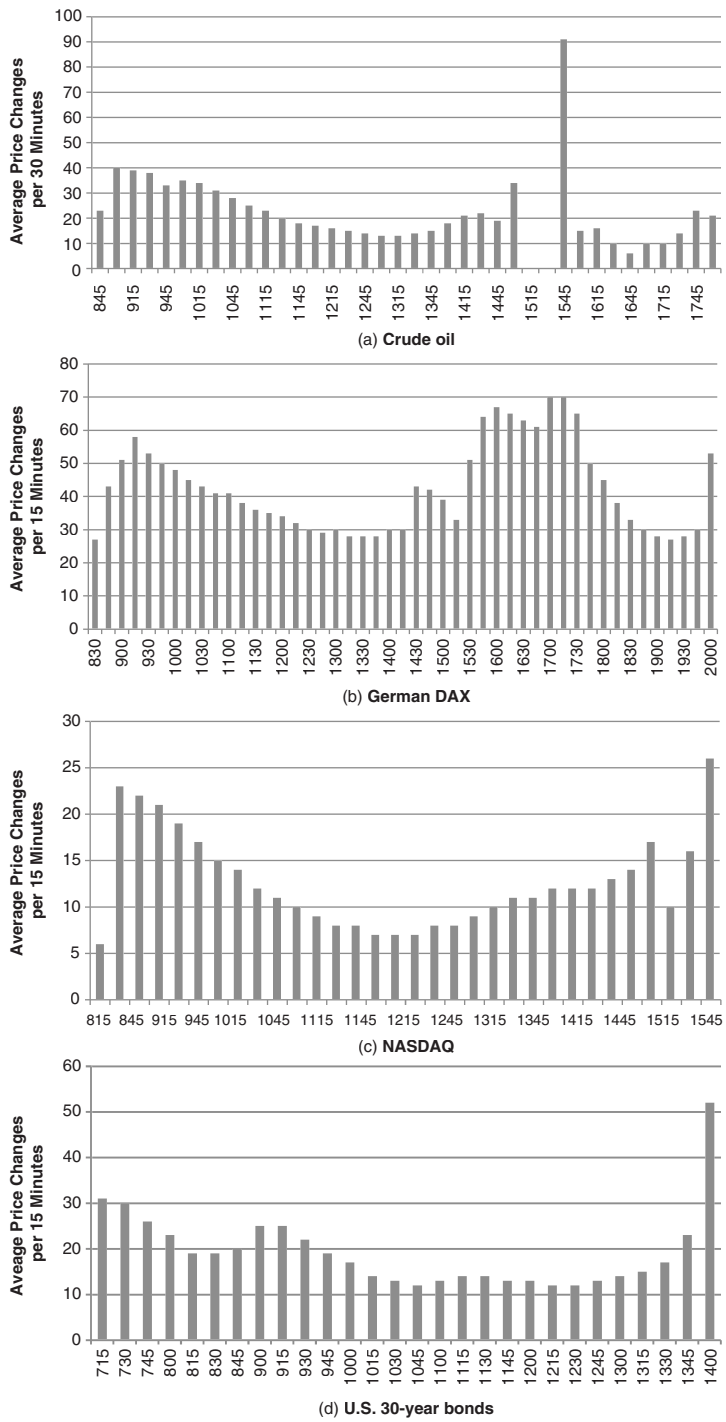
FIGURE 12.15 Intraday, 15-minute tick volume patterns, 1995 to 2005.

will have very low volume and an unclear pattern because we trade it during our business hours, while Japanese business sleeps.

It should not be surprising to see the extreme clustering at both ends of the trading day. Orders enter the market early in reaction to news and events that occurred after the close of the previous day's trading. Many analysts based trading decisions on the previous day's data but perform their calculations and analyses after the close. The end of day is active due to liquidation of the positions of day traders who do not want overnight risk. Orders that allow you to specify that you want to be filled at the settlement have reduced the slippage and increased the benefit of trading on the close.

Relative Changes in Volume

Whether using actual or tick volume, decisions made using, for example, 15-minute volume patterns should compare each 15 minutes of the current day with the same 15 minutes of the prior day or with the average of that 15-minute interval over some range of days. While it may be sensible to use increasing intraday volume as a condition for entering a new trade, if the trade signal is in the second half of the day, then each interval generally increases in volume; therefore, the condition is always true. If the trading signals were in the morning, the condition would always be false.

In order to use intraday volume correctly, you must consider different cases:

1. The volume of the opening bar must be compared to the volume of the same bar on the previous day, or the average volume of that bar on the past n days.

2. The second bar can then show volume as a percentage of the first bar. By keeping a history of this percentage value, we then will know if today's second bar is relatively stronger or weaker than the average.

3. We can track the relative strength or weakness of every bar compared to the opening bar and compared to the average bar of the same time.

Using this information, it is possible to anticipate a stronger or weaker close for the day. It is also generally true that an open on higher volume will indicate higher volume throughout the day. We could then also expect higher volatility. Days with significantly lower volume would be associated with less volatility and less directional movement.

FILTERING LOW VOLUME

Minutes, hours, or even days that have little market activity are likely to be associated with uncertain price direction. If the British pound moves $0.50 in the mid afternoon of the U.S. market, we know that volume is normally light because the London and European markets are closed. The intention of a volume indicator is to identify positive moves

in volume that can be used as a confirmation of price direction. Eliminating those days with low volume, or with marginal price moves, may increase the dependability of the volume indicator.

Removing Low-Volume Periods

Including low-volume periods in a volume index may make the index movements unreliable. There is a similar situation in statistics, where a number of statistical values are slightly skewed to one side. Each value on its own is not noticeably important, but collectively they cause a bias. Some analysts believe that the collective weight creates significance, while others favor ignoring values that are so small they are insignificant and perhaps confusing.

The decision to remove low volume days, or intervals, means that a series of low volume days that results in a collective change will be ignored. If you have decided that removing those days makes sense, then measure the daily or intraday volume against a threshold created using the average volume minus one or two standard deviations of the volume. Using a 1 standard deviation filter will remove the lowest 16% of the days; a 2 standard deviation filter removes 32% of the days.

This type of filter is best applied to a volume index, such as On-Balance Volume. For example, we find that the average volume on the New York Stock Exchange is 1.5 billion shares, and 1 standard deviation of the volume is 0.25 billion shares. We decide that the volume filter is 2 standard deviations; therefore, any day with volume below 1.0 billion shares will be ignored. Applying this to On-Balance Volume, using TradeStation code:

```
Volumethreshold = average(volume, n) – volumefilter*stddev(volume, n)
 If volume < volumethreshold then VFOBV = VFOBV[1]
    else VFOBV = VFOBV + (close – close[1])/
       absvalue(close – close[1])) + volume;
```

we would effectively ignore all days, leaving the OBV unchanged, when the volume was below the threshold.

Removing Volume Associated with Small Price Moves

If you agree with the concept of removing days with low volume, then there is the similar case of very small price moves. Indicators such as On-Balance Volume post all volume as either a positive or negative contribution to the index, based on the direction of prices on that day. It is fair to question the validity of posting all volume to the upside when the S&P 500 closed up a minimum move of 25 basis points (+0.25) or the Dow closed up 1 point. It could just as easily have closed down that amount. In a manner similar to filtering low-volume, periods in which prices moved very little may be eliminated by using a standard deviation of the price changes as a filter. Days that are within ±0.10 or ±0.25

standard deviations of the average would be ignored. The *Price-Filtered On-Balance Volume*, *PFOBV*, would then be found using the code:

```
pricethreshold = average(price - price[1], n) +
    f*stddev(price - price[1], n)
if absvalue(price - price[1]) < price threshold then PFOBV = PFOBV[1]
    else PFOBV = PFOBV[1] + ((close - close[1])/
        (absvalue(close - close[1])) * volume
```

where n = the number of periods in both the average and standard deviation
        f = the number of standard deviations used to filter minimum volume

Note that, in the case of a minimum price threshold, the rules look at price change, which can be positive or negative. For a volume threshold, there is only a one-sided test using the value of volume.

MARKET FACILITATION INDEX

In weighing the likelihood that prices are indicating a direction, rather than a false start, the tick volume compared to the price range for the same period, called the *Market Facilitation Index*,[14] can measure the willingness of the market to move the price. This concept is interesting because it is not clear that high volume results in a large price move, although high volume appears to set up the conditions for high volatility. If the Market Facilitation Index increases, then the market is willing to move the price; therefore trading that benefits from higher volatility should improve.

$$Market\ Facilitation\ Index_t = \frac{Trading\ range_t}{Volume_t} = \frac{H_t - L_t}{V_t}$$

The results of combining the change in tick volume and the Market Facilitation Index are interpreted as:

Tick Volume	Market Facilitation Index	Interpretation
Up	Up	Confirmation of direction
Down	Down	False direction, do not take trade
Down	Up	Poor entry timing, approach with caution
Up	Down	Potential new trend, end of old trend

---

[14] Bill Williams, *Trading Chaos* (New York: John Wiley & Sons, 1995).

Spreads and Arbitrage

For both spreads and arbitrage, traders seek to take advantage of price differences, relative differences, or anticipated differences. Positions taken in opposing directions in related markets, contracts, options, or shares are generally referred to as a *spread* or *straddle*. When a long and short sale are entered simultaneously in two related stocks, such as AMR and UAL, the strategy is called *pairs trading*. When the dynamics of the spread can be definitively calculated, such as the price of two bonds of the same maturity and the same grade, or the price of gold in two different locations, the transaction can be considered an *arbitrage*. (When academics use the term *arbitrage*, it implies risk-free, although nothing is risk-free.) When the reason for two markets to diverge is not a fundamental disparity, such as a location arbitrage (the difference in price caused by transportation between two locations), but rather market forces acting differently on two similar products, the trade is a *relative value arbitrage*.

For futures markets, the most common use of the term *spread* relates to two delivery months of the same market. This can also be called an *intramarket spread*, an *interdelivery spread*, or a *calendar spread*. For example, a trader may take a long position in March Treasury bonds and sell short the June contract (for the same year). The expectation is that the yield curve will steepen (yields on short-term maturities will drop faster than long-term ones). An *intermarket* or *intercommodity spread* can be taken in related futures markets, such as the S&P 500 and the Russell 2000, gold and silver, crude oil (WTI) and Brent North Sea crude, or the U.S. 10-year note and the EuroBund. Because one product or market is not a substitute for another, these are all relative value spreads. Globalization has caused many markets to move together, so when they diverge we have reasonable confidence that they will come back together.

Spreads can be simple or complex, but they generally offer the ability to remove *directional risk*. By being long one market and short another (a short sale in stocks), a strong directional move or a market crisis should move both legs of the spread in the same direction. One leg profits, the other loses, and the net is near zero. This is very different from the typical long-only portfolio in equities or commodities and offers diversification for the investors.

565

DYNAMICS OF FUTURES INTRAMARKET SPREADS

A spread is most often a way to reduce both the risk and, consequently, the potential profit that exists in an outright long or short position. However, futures trading benefits from lower margins for spreads, allowing further leveraging and restoring at least part of that risk. That is not usually the case with stocks, where full financing is required for both the long and the short sale.

An *intramarket spread* is placed in two delivery months of the same market, a combination of futures and physicals in the same market, or in two options of different strike prices or maturities on the same underlying. The reduction in risk depends on the relationship between successive delivery months, which varies considerably across markets. These risks can be summarized as follows:

- *Financial markets* nearest to delivery react faster and with greater magnitude to changing interest rates, economic data, and supply and demand situations. The deferred contracts will respond slower because the lasting effects are not as clear. There are notable exceptions when changes in interest rate policy target longer-term rates and do not take effect immediately.
- *Precious metals* markets, such as gold, platinum, and silver, were once considered *pure carry markets*. Successive delivery months would always trade at a higher price due to the interest rate component of holding the physical product in inventory. If the price of the metal rises or interest rates rise, the cost of carry increases, and the spread between months widens. If metal prices or interest rates decline, the spread narrows. If metal prices and interest rates move in opposite ways, the effect on the spread is dampened to varying degrees. There has been an evolution in the industrial use of precious metals, primarily gold and platinum in electronics, which has increased the consumption and changed the price patterns.
- *Industrial metals*, such as copper, will show normal carrying charges under most circumstances but are affected by demand to the extent that prices have been known to *invert* (*backwardate*) for significant periods of time. Because of its fundamental changes, silver prices inverted for the first time in 1997. The *nonferrous metals*, traded on the London Metals Exchange (LME) are all used in housing and may move together based on the general outlook on the economy.
- *Foreign exchange* rates are dependent on the prevailing economic outlook for the specific country, net of inflation, and often related to their balance of trade. A stable economy will show nearly unchanged forward rates; a weakening economy will cause the deferred contracts to be discounted. Prices tend to become more volatile as they move away from equilibrium, whether higher or lower. Countries prefer to see their exchange rates change slowly. A fast rise in value, which adversely affects exports, may be met with central bank intervention, as has often been seen in the Japanese yen. This makes for a bumpy price pattern. Because exchange rates are quoted in terms of other currencies, prices can only be viewed relative to another economy.

- *Agricultural products (crops)* contain a well-defined carrying charge within each crop year, which begins at the end of harvest. Expected variations in both supply and demand have made delivery month patterns differ from one another; however, prices rarely achieve *full carry*. Spreads are typically entered in the new and old crop deliveries that will show the greatest separation. Increased supply from South America, as well as on-farm storage in North America, has affected the patterns during the winter months in the northern hemisphere, a time when demand and supply have historically been fragile.
- *Livestock* markets are products that cannot be stored and redelivered; therefore, the prices for any one contract month are based on anticipated supply and demand at the time of delivery. Feedlots and farmers have been known to deliver early when prices were high or when production costs were rising; however, this will cause an irreplaceable shortage in the deferred months (Figure 13.1).[1] Patterns in delivery months are a combination of the number of livestock on feed, expectations of marketing, and the price of grain.

CARRYING CHARGES

Carrying charges are an essential part of forward prices. The *carrying charge* structure of a market determines the underlying differences in the relative price of each contract and the price pattern of the delivery months. It is also called the *term structure*. For metals and most storable commodities, these charges consist of interest, storage costs, and insurance. All else being equal, a market with *normal carry* can be expected to show steadily higher prices for deferred contracts based entirely on the cost of carry. Prices that move out of line, where the nearby is unusually inexpensive, may be corrected by taking delivery of the nearby and redelivering against the next contract.

It is always possible for changes in supply and demand to alter the relationship of futures contracts so that the carrying charge pattern no longer appears normal. Although this is not likely to occur in the metals, it often happens that an immediate demand for food can cause short-term increases in prices without substantial effect on more deferred months. This situation is called an *inverted market*, *negative carry*, or *backwardation*, even though the carrying charges are implicit in the price regardless of the pattern.

In the financial markets, a similar situation occurs when the demand for money increases dramatically in the short term. If the market expects that the demand will be short-lived, the rates on short-term maturities are significantly affected while the longer-term rates may show only a minor change.

In the metals markets, a normal carry is called *contango* (a term coined by the London Metals Exchange when referring to copper), and a negative carry is *backwardation*.

---

[1] Perry J. Kaufman, "Technical Analysis," in Nancy H. Rothstein, ed., *The Handbook of Financial Futures* (New York: McGraw-Hill, 1984).
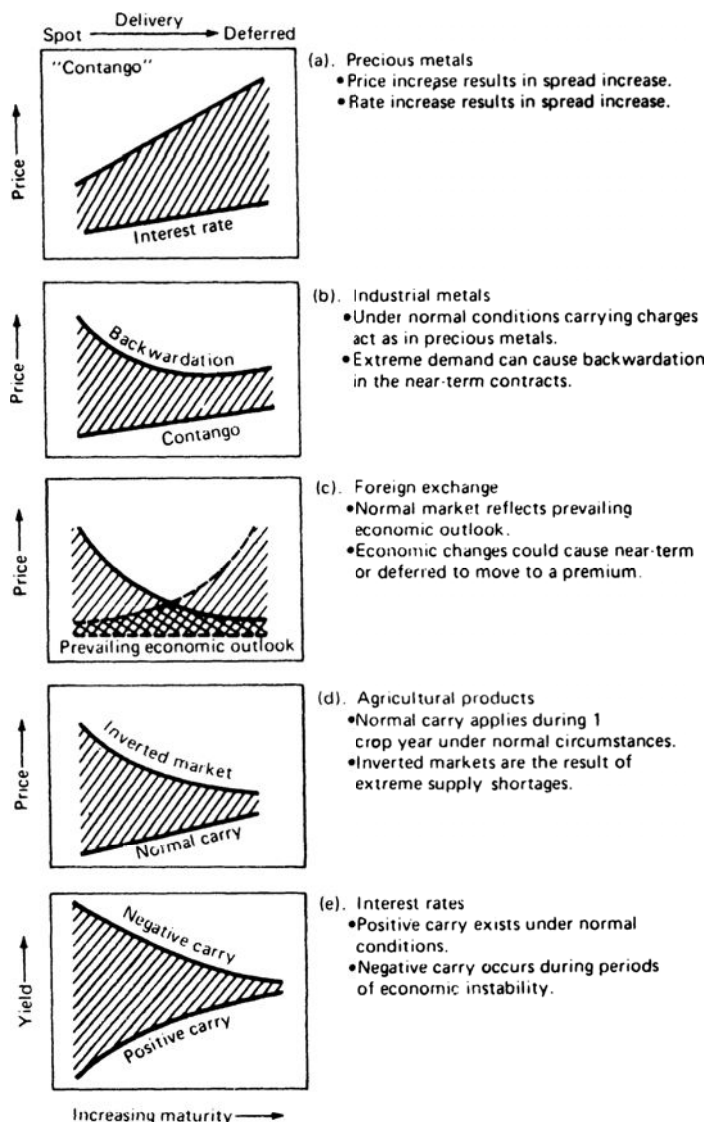
FIGURE 13.1 Interdelivery price relationship and terminology.
(a) Precious metals. (b) Industrial metals. (c) Foreign exchange.
(d) Agricultural products. (e) Interest rates.

Changes in supply and demand that lead to backwardation are most likely to occur in the industrial metals where the commodity is consumed at a greater rate. Gold, silver, and platinum, the *precious metals* markets, have traditionally had a uniform price relationship in the deferred delivery months. Any deferred delivery price can be calculated by adding the current rate of interest (plus storage and insurance) to the value in the cash market. If interest rates rise and the price of gold remains the same, the carrying charges

and the spread will increase. If the price of gold rises but interest rates remain constant, the carrying charges and spread will also increase because the increased value of the contract results in higher interest costs. Similarly, lower carrying charges and a narrowing spread will occur if either or both the rate of interest or the price of gold declines. Precious metals are subject to *implied interest rate* spreading, discussed in the next section.

The terms used in referring to carrying charges in the financial markets are the same as those just mentioned, but the concepts are different. *Carry* is a term describing the yield-curve relationship. The concept of positive or normal carry is a curve that increases in yield with the time of maturity. The longer an investor is willing to commit money, the greater the yield. Negative carry can also exist for short periods of time. When economic conditions become unstable, the interest on short-term investments may increase sharply although longer-term rates will increase only slightly. Investors anticipate a correction in these short-term distortions, and do not often move money committed for longer periods for fear that rates will return to positive carry.

SPREADS IN STOCKS

In the stock market, spreads are actively traded in related companies or grouped into baskets.

- *Pairs trading*, which is buying one stock and selling short another in the same business, such as Verizon and AT&T, will profit from the competitive advantage of one firm over the other while removing the directional risk of the overall market.
- *Market neutral trading* can be thought of as a basket of stock pairs or a basket of long positions traded against a basket of short sales in the same industrial group or in multiple industrial groups. It may also be a basket of stocks that are ranked as "outperform" against a basket of those expected to "underperform." Generally the stronger stocks are sold short and the weaker bought; however, the opposite may be done in a trending market.
- *Dividends versus Treasury yields* are an important benchmark for investors, who are always seeking the highest return. Most public utilities are in direct competition for low-risk investments targeting interest rate income. When Treasury yields become extremely low, as they did in early 2003, investors are willing to take greater risk to receive higher returns and begin looking for stable companies with high-dividend returns. When Treasury yields are very high, they draw investments away from many other vehicles, including the stock market.
- *Cross-market spreads* will take one position in a stock that is highly dependent upon a physical commodity, such as a precious metal or oil, and the other in the futures market for that commodity.

Globalization

Globalization has increased the correlation in nearly all world markets, whether financials or commodities. Countries all compete for investment dollars (or euros, yen, and now the Chinese yuan). Higher interest rates and a strong economy in the United States would put pressure on Europe to raise its rates or suffer a shift of funds away from the euro into the dollar. A weak U.S. economy has a ripple effect on many of the world economies because the United States is an extremely large importer. In general, money flows to those countries with the highest yield on sovereign debt, net of inflation and political instability. In a crisis, such as the 2008 subprime meltdown, money will flow out of all markets into cash, causing correlations in all assets to approach 1.0.

Technology facilitates globalization. European, U.S., or Asian economic reports are disseminated instantly. When the U.S. government releases its reports, often at 8:30 A.M. Eastern Time, the world knows those statistics within seconds. It is fair to say that the world reacts at the same time to these reports. Electronic trading, which allows 24-hour access in many markets, provides the means for price discovery and speculation.

Globalization has greatly increased the number of opportunities for spreading markets. We can expect the European interest rates to fall if the U.S. lowers its rates. We can expect the U.S. dollar/euro (USDEUR normally quoted as EURUSD) exchange rate to fall if European economic growth is stronger than U.S. growth, or if the balance of trade between the countries shifts. The meetings of the *Group of 7* ("G7"), or G9, or G18, further increase the chance of a unified policy by setting common expectations for interest rates and exchange rates among the largest economies. While there is a clear dependence of one market on the other, the effects are highly variable. Because these relationships continue to change, there is great opportunity in spreads that span world markets. Recent high levels of debt in developed countries, and their method of reducing that debt, have caused rapid changes in FX crossrates.

Spreads are often unique to a specific market situation and cannot be generalized. The trader must first understand the basis of the spread relationship before any strategy can be formulated. It is most important that a trader understand the conditions under which a spread, or even an arbitrage, will fail. This chapter presents many approaches to spreading that are very specific; examples are equally limited in scope and application.

SPREAD AND ARBITRAGE RELATIONSHIPS

The relationship between two markets, or between various deliveries and trading vehicles within the same market, will determine the type of trading strategy that may be applied. The types of spreads and arbitrage situations that are most often watched are:

1. *Substitute products*, such as wheat and corn or cattle and hogs. Product substitution ranges from those markets that are nearly identical (e.g., 3-month T-bills and 3-month Eurodollars), to cotton and soybeans, which share the same land and growing season.

2. *Location spreads*, including gold in New York, Chicago, and London, cocoa and heating oil (gasoil) in New York and London. The price at one location should never be greater than the cost of moving the product from another location.

3. *Carrying charge spreads* and *cash and carry*, where one delivery month is out of line with others based on the cost of storing the physical product and redelivering it at a later date.

4. *Product relationships*, such as crude oil versus heating oil and gasoline, and soybeans versus soybean meal and oil. Costs for converting the raw material to its component products define relative prices.

5. *Usage spreads*, including the hog-corn ratio, feeder cattle-corn-fat cattle, cocoa-sugar, broilers-corn, and lumber-plywood. If the cost of corn increases, then the cost of livestock increases. If the market will not absorb the higher end-product cost, as determined by deferred futures prices, then livestock is brought to market early.

6. *Pure mispricing in financial markets*, such as different exchange rates offered by different banks around the world, and interest rates of the same maturity and the same grade, where there is no actual cost of delivery or carrying charges.

7. *Interest rate strips*, where pieces of varying interest rate maturities are put end-to-end and must be equivalent to the comparable rate offered on longer-term maturity.

RISK REDUCTION IN SPREADS

As long as there is some fundamental relationship between the two markets being spread, A and B, the resulting spread return stream will be less volatile than the average volatility of the two spread components. In mathematical terms, the variance of the spread is

$$Std^2_{A,B} = Std^2_A + Std^2_B - 2 \times Cov_{A,B}$$

where Std = the standard deviation
Cov = the covariance

When the two components of the spread are more closely related, they vary in the same way at the same time, and the covariance increases. This in turn reduces the variance of the spread. Covariance is calculated by first finding the differences between the corresponding prices and their average values, multiplying them, and finding the average over period N.

$$Cov_{A,B} = \frac{1}{N} \sum_{i=1}^{N} (A_i - avg_A)(B_i - avg_B)$$

It is easy to compare the risk of the spread with the risk of the individual markets using a spreadsheet. Put the prices of market A in column A, the prices of market B

in column *B*, and the spread difference or ratio in column *C*. Put the returns, or daily changes, of *A*, *B*, and *C* into columns *D*, *E*, and *F*. Calculate the standard deviations for all of the values in each of columns *D*, *E*, and *F*, and simply compare them. The stronger the relationship between *A* and *B*, the smaller the standard deviation of *F*.

ARBITRAGE

When the two legs of a spread are highly correlated, the trade is an *arbitrage*. When it can be executed often and corrects within minutes, or even milliseconds, it is also *high frequency trading*. More complex arbitrage can take advantage of price differences in related products and last days or weeks; however, absolute risk is proportional to the time in the trade. True arbitrage has, theoretically, no trading risk; however, those opportunities are also associated with small profits and low liquidity. For example, a *spatial arbitrageur* using the Interbank market might call (or access electronically) one bank in Tokyo and another in Frankfurt to find their rates on the Mexican peso. If they differ, the trader would buy the peso from one bank and sell the peso at another provided:

- The price difference was greater than the bid-asked spread, the implied cost of execution.
- The arbitrageur has proper credit established with both banks.
- The transaction can be performed *simultaneously* (by telephone). This requires one trader with a telephone in each ear or two traders working side by side.
- If both banks make a market electronically (which is preferable), then arbitrage of rates between banks can be done with direct order entry systems.

Large-scale arbitrage has become the domain of major financial institutions that employ many traders, each provided with high-tech computer displays, sophisticated analytic software, and lots of telephones. These traders specialize in specific interest-rate markets, foreign exchange, individual stock selection, or less often, precious metals. They constantly scan quotes from across the world to find price differences, then act quickly using cash, forward, futures markets, and options. They trade large quantities to profit from small variations. For the interest rate markets, there are computer programs that compare the various types of coupons and maturities in order to identify an opportunity quickly. Such operations have become an integral part of the financial industry; they keep prices in-line around the world and generate steady profits.

Pricing of Futures Contracts

The relationship of one futures market delivery month to the spot price of that market is different according to the type of product, which has already been

described in general terms. The mathematics of some of these relationships can become very complex, and the reader is referred to texts that deal specifically with these subjects.[2] The following sections describe the most important features of these relationships.

Convenience Yield

Convenience yield is the forward price of a commodity less the cost of storage. It is the benefit of holding the underlying physical product compared to a futures, forward, or options contract. For example, the price of December gold futures, F, on June 1 is $1400, and the current spot price, S, is $1350. The borrowing rate, r, is 5.0% and there is no cost of storage. Then the convenience yield, CY, is

$$F_{t,T} = S_t + (S_t \times r_T \times (T-t)/365 + CY_{t,T})$$

Rearranging this to solve for the convenience yield,

$$CY_{t,T} = F_{t,T} - S_t - (S_t \times r_T \times (T-t))/365$$

Substituting the values in the example,

$$CY_{t,T} = 1400 - 1350 - \left(1350 \times 0.05 \times \frac{(182.5 - 0)}{365}\right) = 16.25$$

Then 16.25/1350 gives a 6-month return (convenience yield) of 1.20%. Had the interest rate been lower, all of that difference would have gone into the convenience yield.

Using the same concept, the convenience yield between any two futures markets can be calculated. Replace the spot price with the nearest futures, the futures price with a more deferred futures price, and $T - t$ becomes the number of days between the two futures contract deliveries. Geman gives the relationship between the spot and futures price as

$$F_t^T = S_t \times e^{(r - CY)(T - t)}$$

Arbitraging the Carry of Storable Commodities

Storable commodities can be purchased in the cash market, stored, and sold at a later time. Storable commodities range from gold to grain, from crude oil to orange juice. They can be delivered at the maturity of a futures contract, held in storage, and redelivered against another deferred contract. This puts an upper limit on the amount of the carrying

---

[2] An excellent source is Heylette Geman, *Commodities and Commodity Derivatives* (Hoboken, NJ: John Wiley & Sons, 2005). Also, Aswath Damodaran, *Investment Valuation* (New York: John Wiley & Sons, 1996), 448–458, which provides the basis for some of these formulas. The first popular source book was Marsha Stigum, *Money Market Calculations* (Chicago: Dow Jones-Irwin, 1981).

charges that can be implied in futures prices. The difference in cost between holding the physical commodity and buying it on the futures market are:

- The *financing cost* involved in the purchase of the physical commodity:

$$Cost\ of\ interest = S_t \times \left((1+r)^{\frac{T-t}{365}} - 1 \right)$$

- The *cost of storage*, if any.
- A *convenience cost* for not buying the physical product, and the ability to be able to sell it at any time.

These three costs are added to the spot price to get the fair value of the futures price at the time of delivery. However, this relationship represents the ideal case. If you add more realistic features, including a separate borrowing and lending rates, R_b and R_a, where $R_b > R_a$, and assume that the short seller cannot recover the saved storage costs, k, and must pay transaction costs, c_s, as well, you get a normal range in which futures prices can fluctuate:

$$(S_t - c_s)(1 + R_a)^T < F_t < S_t \times \left((1 + R_b)^T + kT \right)$$

where T is the time to delivery of the futures contract or the holding period, in years. When futures prices move outside this range, there is an opportunity for arbitrage.

Interest Rate Parity

One well-known, second-order arbitrage combines foreign exchange forward rates with interest rate parity. Consider a U.S. corporation that would like to invest $1 million for the next 6 months. The current U.S. T-bill return for the next 6 months is lower than the comparable rate in Europe, and the inflation rate is about the same. The corporation is faced with the decision of whether to convert U.S. dollars to euros and invest in European time deposits or accept the lower U.S. rates. The decision is made easier if the corporation purchases goods from Germany, since it must eventually convert U.S. dollars to euros to satisfy payments; the conversion cost will then exist with either choice.

What if the value of the euro loses 1% against the U.S. dollar during the 6-month investment period? A corporation whose deferred payment is stated in euros suffers a 1% loss in the total amount received. If the corporation bought 6-month euro-denominated 100,000 CDs with a return of 4%, it now repatriates €102,000, or $100,980, a net return of only 98 basis points, or an annualized rate of 1.96%. A U.S. trader buying and selling futures contracts denominated in euros, yen, or reals, needs to account for the changes in the exchange rate of the invested capital (usually margin). The *interest rate parity theorem* will normally explain the differences between the foreign exchange rates and the relative interest rates of countries. It states that *the forward rate of a currency is equal to its present value plus the interest earned in that country for the period of the*

forward rate.[3] Using the futures or Interbank market for the 1-year forward exchange rate (EUR_{1yr}), the spot rate for the current value (EUR_{spot}), and the annual interest rate in the European Community (I_{EC}), we get the relationship

$$EUR_{1yr} = EUR_{spot}(1 + I_{EC})$$

The forward value of the U.S. dollar can be expressed similarly as

$$USD_{1yr} = USD_{spot}(1 + I_{US})$$

Compute the *implied interest-rate/forward-rate parity* by dividing the first equation, EUR_{1yr}, by the second, USD_{1yr},

$$EURUSD_{1yr} = EURUSD_{spot} \times \frac{1 + I_{EUR}}{1 + I_{USD}}$$

Naming conventions in foreign exchange (FX) interprets EURUSD as the number of U.S. dollars per euro. For example, if U.S. interest rates are 4%, European rates are 6%, and the current exchange rate $EURUSD_{spot}$ is 1.450, the 1-year forward rate would be

$$EURUSD_{1yr} = 1.450 \times \frac{1.06}{1.04} = 1.4779$$

The forward exchange rate of 1.4779, higher than the current rate, is entirely based on the rationale that higher yields in the EC will attract more investment and further weaken the U.S. dollar. Any difference between this calculation and the actual forward exchange rate available in the market is based on anticipation of other factors.

Backwardation in exchange rates can be described as downward-biased estimates of future spot prices. If we consider the case of the EURUSD in the previous example, and the forward interest rate of the euro is less than the comparable forward rate in the United States, then the ratio $\frac{1 + I_{EC}}{1 + I_{US}} < 1$ and the forward rate of the EURUSD will be lower than the spot rate.

When testing strategies using foreign exchange (crossrates), it is not good to use the spot rate and then trade futures. The interest rate parity will cause prices to move systematically for you or against you, but your results using forward prices will be different from those using cash prices. A program, *TSMInterestRateParity*, can be found on the Companion Website.

Crossrate Matrix Evaluation

Given the large number of currency exchange rates that can be quoted in terms of any other currency, there are many combinations of rates that might provide an opportunity for arbitrage. The process for uncovering those inconsistencies that offer opportunity is easier if you put all of the rates into a table format often seen in the newspapers (see Table 13.1).

---

[3] James E. Higgens and Allen M. Loosigian, "Foreign Exchange Futures," in Perry J. Kaufman , ed., *Handbook of Futures Markets* (New York: John Wiley & Sons, 1984).

TABLE 13.1   Major Crossrates as of April 29, 2011

	AUD	CAD	CHF	CNY	EUR	GBP	JPY	SGD	USD
AUD		1.036955	0.94768	7.101643	0.736818	0.655798	88.87316	1.339591	1.09374
CAD	0.964362		0.913906	6.848553	0.710559	0.632426	85.70588	1.29185	1.054761
CHF	1.055209	1.094204		7.493716	0.777496	0.692004	93.77973	1.413548	1.154124
CNY	0.140812	0.146016	0.133445		0.103753	0.092345	12.51445	0.188631	0.154012
EUR	1.357188	1.407343	1.28618	9.638264		0.890041	120.6176	1.818077	1.484411
GBP	1.52486	1.581212	1.445079	10.82901	1.123544		135.5192	2.042689	1.667801
JPY	0.011252	0.011668	0.010663	0.079908	0.008291	0.007379		0.015073	0.012307
SGD	0.746497	0.774084	0.70744	5.301352	0.550032	0.489551	66.34351		0.816473
USD	0.914294	0.948082	0.866458	6.49299	0.673668	0.599592	81.2562	1.22478	

Reading down the last column of Table 13.1, *USD*, the numbers represent the amount of U.S. dollars needed to buy 1 Australian dollar (1.09374), 1 Canadian dollar (1.054761), and so on. If we call the table of crossrates C, then every crossrate in C is related to every other crossrate according to:

$$a_{ij} \times a_{jk} = a_{ik}, \text{ for all } i, j, k$$

where the first subscript is the row, the second is the column, and each value of a represents an element in the table, with a_{ji} representing currency i in terms of currency j. Therefore, values in the table are expressed as column-row rates, such as USDGBP column 9, row 6, or EURJPY in column 5, row 7. The top right triangle of the table is exactly the inverse of the lower left part, and can be calculated as:

$$a_{ij} = 1/a_{ji}$$

Any two rates in Table C can be multiplied together to get a third rate,

$$a_{ij} \times a_{mn} = a_{in}$$

As an example in Table C, multiply the EURUSD (a_{59}) $\times$ USDJPY (a_{97}) = EURJPY (a_{57}). Then $a_{59}(1.484411) \times a_{97}(81.2562) = a_{57}(120.6176)$, proving that the crossrates in Table C are correct.

The difference between the implied interest-rate/forward-rate parity and the current forward prices may offer arbitrage opportunities. Using the same crossrates in Table 13.1, we can find the forward rates for June 2011 delivery and the current 3-month Libor rate using the interest rate parity formulas given in the previous section, then we can calculate our own forward price and compare this against the actual futures price. The difference is the *opportunity differential*.

Program Trading

Arbitrage is most often associated with institutional trading. Because profits on individual positions are small, it is necessary to trade large numbers; hence, the capital

requirements can be equally large. The most well-known of all arbitrages is called *program trading*, the process by which a stock index futures market and the cash index (the weighted average of the actual stock prices) are kept in the proper relationship to one another.

There is nothing secret about program trading. The premise is that cash and futures prices will come together (to within a small price difference) by the time the futures contract expires. If prices converge sooner, profits are larger. At expiration the buyers and sellers must settle in cash at the current cash market value. If the difference between the cash and futures price exceeds the cost of trading plus a profit, then the two markets can be arbitraged. A profit in the trade, P, is

$$P = |C - F| - C \times r \times t - E$$

where C = the cash price of the stocks (the cash index)
   F = the price of the comparable futures index
   r = the annualized interest rate
   t = the time to maturity in years
   E = the execution cost

For a successful trade, P must be greater than some value that makes it worth trading. The arbitrageur must also have a fully automated order-entry system so that orders to buy all of the stocks that comprise the index, and sell short the corresponding futures market, can be done instantly.

Program trading may also be triggered when futures prices differ from fair value by a sufficient amount to be profitable. The *fair value* is the price at which cash and futures will converge at delivery based on today's values.

$$\textit{Fair value (FV)} = C \times (1 + (r - D) \times t)$$

where D = the annualized dividend payment

In addition to accepting a small differential between cash and futures at the time of delivery, you must know:

- The transaction costs of executing both the stock and futures side of the trade. This includes both commissions and an estimated slippage for each of the stock issues that must be bought.
- The need to deal with an *uptick rule*. If an uptick rule is currently enforced, then the trade is much more practical when futures are above cash so that futures are sold and stocks are bought. The uptick rule requires that stocks only be sold on an uptick, which makes it impossible to guarantee that short sale positions can be entered in each stock, the amount that can be sold, or that the trade can be completed in a reasonable time. It can be done, but it is very difficult and may require some imaginative variations.

• The availability of all stocks (or a designated subset) for short sales. While many stocks can be sold short without notification, large institutions may keep a passive portfolio of stocks representing a particular index so that they can be sold as part of a program trade.

Program trading can work for any index futures, which includes the 30 Dow Industrials, the S&P 500, the Russell 2000, small caps, and even sectors, trading as ETFs. Trading the larger index futures, such as the S&P 500, involves large cost for setting a position in all of the 500 stocks that comprise the Index. There is also competition from other institutions. This limits the profit potential of the trade to the smallest acceptable amount over the break-even level.

Program trading is separated into buying programs and selling programs. A *buying program*, the easier of the two, buys stock and sells futures when the difference is profitable. A *selling program* sells stocks and buys futures, a difficult implementation because of the possibility of an uptick rule for U.S. stocks and the availability of stocks for short sales. Adding more realistic assumptions of distinct borrowing and lending rates, r_b and r_s, made in the basic futures arbitrage discussed earlier, plus separate execution costs for buying stock E_b and selling short E_s, the *fair value at delivery* can vary within the band:

$$(C - E_s)[1 + (r_b - D) \times t] < FV < (C + E_b)[1 + (r_s - D) \times t]$$

Those readers interested in examples can find a *program trading calculator* online. This allows you to enter the cash index value, the buying interest rate and the selling rebate, the dividend return, and the time to maturity. You get the levels at which an arbitrage can be set. Not included are the transaction costs.

Finding a Representative Subset of an Index

An arbitrage would be much easier and less costly if there were a smaller group of stocks or futures markets that performed the same as the entire index. If the index is weighted by capitalization, volatility, or some other characteristic, then sorting the components of the index in order of largest capitalization or highest volatility could allow you to find a reasonable subset. In terms of capitalization, it is likely that a smaller group of stocks would represent 90% of the index value; therefore that subset would perform in a way very similar to the index. While the S&P 500 is comprised of companies that all have relatively high capitalization, many of the index markets for countries with smaller economies will be dominated by a few large companies.

Another approach to finding a subset of markets that closely tracks the index value is by using stepwise regression, inputting all of the markets that compose the index, and examining the weighting factors assigned to each in the answer. Those markets with the highest weighting factors are the most significant in the movement of the overall index. The regression would need to be rerun frequently, and the basket adjusted, to stay aligned with the index. The natural trade-off for using a subset is that there is added risk based on how the subset tracks the full index.

Fair Value

If we could know the correct price, the *fair value*, of a stock or futures market, we would automatically know what trade to enter. For the S&P futures, the theoretical fair value can be calculated from the cash index (SPX),[4]

$$Theoretical\ Futures\ Fair\ Value = SPX \times [1 + r \times (e/360)] - d$$

where r = risk-free interest rate
     d = aggregate dividend rate of the stocks in the SPX
     e = number of days until expiration of the futures contract

This is similar to the calculation for single stock futures in the section that follows. The difference between the calculated fair value and the futures contract value is an opportunity for arbitrage; however, trading the cash market is not usually possible; therefore, substituting the ETF SPY, or finding the difference between the first and second futures deliveries, should be more practical.

Merger Arbitrage

Merger arbitrage, also called *risk arbitrage*, tries to take advantage of mergers and acquisitions between two public companies. It can be an important part of a financial institution's portfolio. There are two main strategies for merger arb:

1. Find a company that is a candidate for acquisition based on market share, recent earnings (usually not doing well), providing a missing product line for a major player in the same industry, or a company that can swing the balance of market share to more than one larger company if it is acquired. (Recently, T-Mobile was targeted as such a company by the larger wireless providers.) Buy a call option, or some combination of options that put you net long, and wait for an announcement. Close out the positions after the announcement or if the potential for acquisition disappears.

This method depends on the research of the trader or analyst, but it is a common industry practice to speculate on takeovers. If right, you are in well ahead of everyone else.

2. More common is a strategy that waits for the official announcement by the board of directors that they have accepted an offer by another company to be bought at, say, a 30% premium over the close of trading on the previous day. If the price is $30 then the takeover price is $40. On hearing the announcement, traders will buy the target stock (the one being taken over) at nearly any price under $40, expecting to get $40 when the merger is finalized. Some traders will take an opposite position, a short sale in the acquiring company, as a market neutral position. However, there is nothing to show that shorting the acquiring company actually reduces the risk of the trade.

---

[4] Adapted from Randy Frederick, "Trading with Fair Value," *Futures* (April 2006).

On the day of the announcement, based on the confidence that the deal will ultimately close, the price of the stock will move up from $30. If there is great confidence, then the stock may trade as high as $39. If low confidence, it may only go to $34. The high confidence trades are the most reliable. If the deal closes in four months, which is typical, then a $1 profit on a $39 investment is a 7.69% annualized return.

The risk in this merger-arb trade is that the deal fails, often based on undisclosed losses or other accounting issues in the target company. When announced, the target stock falls back to its original price, and sometimes lower. This can be a large loss, but it should occur infrequently. Hence, merger arb is considered to have a *short-options profile*—many small profits and a few large losses.

Offsetting the bad events when a deal fails are the cases where another company offers a higher price, say $43. Now the return becomes $4 on $39, or more than 30%. In a portfolio of many deals, these upside surprises go a long way toward offsetting the losses from failed deals.

The most difficult part of trading a merger-arb program is the proper use of investment funds. Given a fixed investment allocated to merger arb deals, it is important to use all the money. If any of it sits idle, then the return on investment is reduced. The difficulty is that no one knows when the next deal will occur, whether it will be big or small (addressing liquidity), or how long it will last. The best returns are when there are many deals in the portfolio, so that one failed deal is only a small part of the total. This will be discussed in Chapter 24, Diversification and Portfolio Allocation.

Single Stock Futures and the Underlying Stock

Single stock futures allow you to buy or sell short a stock for future delivery or cash settlement. The advantages of single stock futures is the lower cost of holding a position, the ability to go short without the restriction of the *uptick rule*, which requires that all short sales (new short positions) be initiated on an uptick, and increased leverage.

A natural arbitrage exists between single stock futures and the cash price of the stock traded on the NYSE, NASDAQ, or another exchange. The decision to enter this spread is the same as for program trading. The price discrepancy must be large enough to allow a profitable trade. For example, if the cash price of IBM is $120, interest rates are 3%, and IBM's dividend yield is 1%, it would cost the trader $120 × 0.03 or $3.60 in lost interest to purchase the stock.[5] The trader also expects to receive $1.20 from dividends over the life of the trade ($120 × 0.01). Ignoring interest on the dividend payment, the price of the futures contract should be

$$Single\ stock\ futures\ prices = Stock\ price \times e^{rt} - \sum_{i=1}^{N} Dividend \times e^{r_i t_i}$$

---

[5] Kenneth Kapner and Robert McDonough, "Doing Your Homework on Individual Equity Futures," *Futures* (March 2002).

where r = the interest rate
            r_i = the forward interest rate over the partial year t

The price of IBM futures should be $120 + $3.60 − $120 = $122.40; therefore, if the difference between the cash price of IBM and the single stock futures price varies because of strong selling concentrated in the cash market, you can buy the cash and sell futures until the fair value is restored. For example, if IBM is trading at $119 on the NYSE and the nearest futures is at $123, you can buy 100 shares and sell 1 futures contract (equal to 100 shares). If the markets settle at $120 and $122.40, respectively, you will have a gain of [($120.00 − 119.00) + ($123.00 − $122.40)] × 100 = (1 + 0.60) × 100 = $160, less commissions.

When trading in a foreign market, for example, a U.S. investor trading the Nikkei 225 Index, the theoretical value of the futures price in T years (where 3 months would be $T = 0.25$) would be

$$Futures\ price = (1 + r_J - d_J)T \times Cash\ price$$

where r_J = the short-term interest rate in Japan
            d_J = the average dividend yield on the Nikkei, in decimal.

Intermarket Index Spreads

The propagation of financial and stock index futures markets, ETFs, sector funds, large and small cap funds, Index funds, and many other structured products have been rampant in the past few years. Every country seems to have introduced its equity index into the marketplace, all with the idea of competing for their share of hedgers and speculators. Combined with the ability to trade electronically from anywhere, there are times when many of these markets move together, offering spread opportunities.

Spreading two stock indexes expresses a particular market opinion. If you consider four popular U.S. indexes—the DJIA, S&P 500, NYSE Index, and the Russell 2000—we see that each represents a different profile of the business economy. The DJIA, limited to 30 very large companies, and weighted by price, is viewed as being the safe haven of stocks. When the economy is not robust, investors move to these very strong companies, which traditionally pay competitive dividends. The S&P 500 is a broader view of high-cap stocks, offering more diversification but still viewed as safe. The NYSE Index is the average of all shares and expresses the broadest view of the economic health. If you believe that the economy is headed for a downturn, but do not want to be net short the entire market, then a spread that is long the Dow and short the NYSE Index would be a lower-risk approach to implementing that position. The Russell 2000 is a small-cap index, representing the weakest part of the market during a poor economy but often the first to recover as a bear market comes to an end.

Figure 13.2 shows the four index futures from August 2002 through September 2003, the end of the bear market that began in January 2000. It is very difficult to see significant differences in any of the indexes without creating a ratio chart, shown in Figure 13.3.
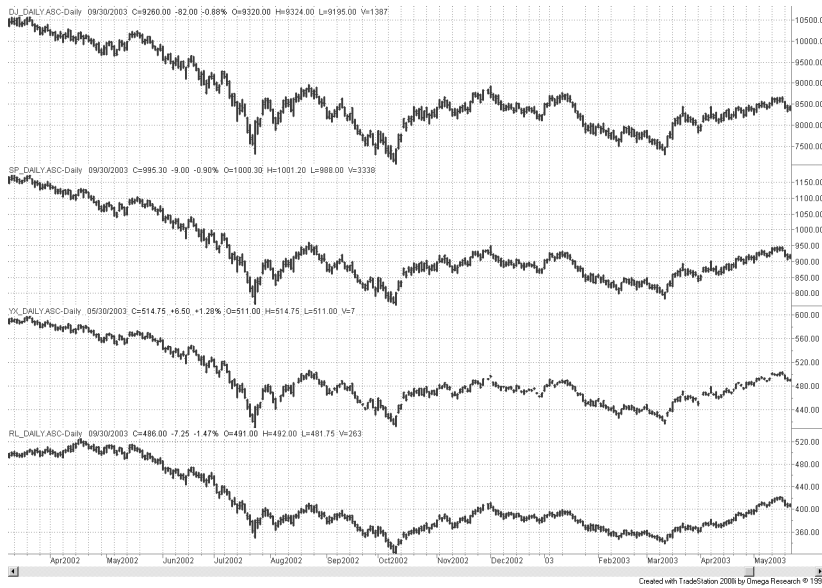
FIGURE 13.2 Major Index futures markets. From top to bottom: Dow Industrials (DJ), S&P 500 (SP), NYSE Index (YX), and Russell 2000 (RL).

A *ratio chart* divides the price of one index by another. In Figure 13.3a the ratio of the S&P 500 and the DJIA varies from 0.104 to 0.110, with a midpoint of 0.107, a range of about 6%. Considering the very different ways that these index values are calculated, this is a narrow range. Of course, all of the Dow components are among the largest in the S&P 500; therefore, there should be a similarity between the two. In the short term, differences in these index markets are mean reverting; that is, they would be traded by
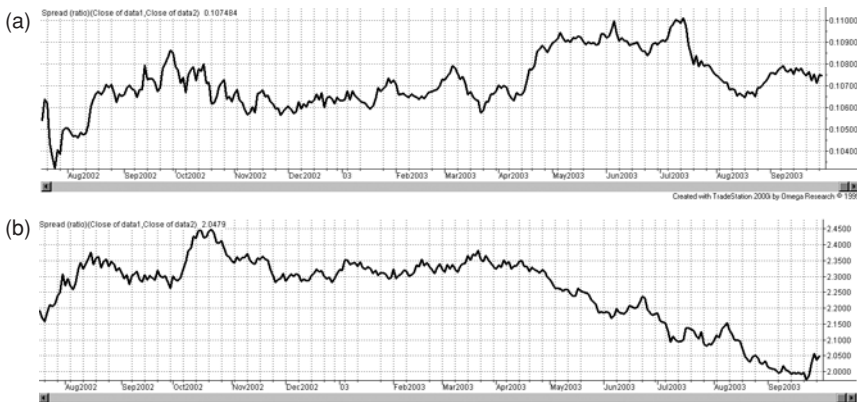
FIGURE 13.3 (a) Ratio of S&P 500 to Dow Industrial. (b) Ratio of S&P 500 to Russell 2000.

buying the lower and selling the higher, expecting prices to return to the midpoint. However, there can be a longer-term trend in the ratio.

The S&P 500 and the Russell 2000 represent different ends of the stock market, the high caps and the small caps. In Figure 13.3b the declining ratio of the S&P compared to the Russell shows that the Russell has been much stronger since April 2003. It is a sign that investors are comfortable with the market and will put their money in stocks that have higher risk. This trend in the ratio would normally be traded by selling the S&P and buying the Russell at a time when an economic recovery was expected.

Pairs Trading

Pairs trading is one of the oldest and most well-established equity trading strategies. It also goes by the names *statistical arbitrage* (*stat-arb*) and *relative value arbitrage*. It takes advantage of price differences between two related stocks, indexes, or ETFs. In its most advanced form, it can be the basis for high frequency trading, where fully automatic systems enter and exit the markets in microseconds. A great deal of money has been made with this method.

The first step is to find two candidate markets. Quantitative analysts have a number of different methods they use to identify good pairs, the two most common are:

1. *Correlations* between 0.5 and 0.7. Correlations tell you about the short-term relationship between two series. In this range, the two series are often moving in the same direction on the same day, but not always. It is the simplest and most popular way to identify pairs.

2. *Cointegration* identifies whether the long-term trend of the two series is the same. If it is, then variations around the trend are trading opportunities.

Cointegration is the better measure.[6] It can be thought of as walking two dogs each on their own lead. They both go forward, but their movements left and right seem random. Cointegration would identify many more pairs than correlation.

Perhaps the easiest and most practical way is to find two companies in the same or a related business and compare the charts. If they are moving the same way most of the time, and react the same direction during a crisis, they are good candidates.

There are numerous ways to identify the entry and exit timing for a pairs trading, but this section will discuss only one.[7] Readers with deeper interest can find more material online or in a number of other books. This approach will use relative value arbitrage, finding two related stocks where one is relatively strong and the other relatively weak. The strategies need to identify when these price discrepancies occur, calculate the position size for each leg, then decide when to enter.

---

[6] Cointegration is a more complex process. Most common techniques are the Engle-Granger 2-step procedure, the Johansen procedure, and the most popular, the Phillips-Ouliaris Cointegration Test. Specialized statistical software is needed. An excellent discussion of cointegration can be found in Carol Alexander's books, in particular *Market Models* (New York: John Wiley & Sons, 2001).

[7] See Perry Kaufman, *Alpha Trading* (Hoboken, NJ: John Wiley & Sons, 2011).
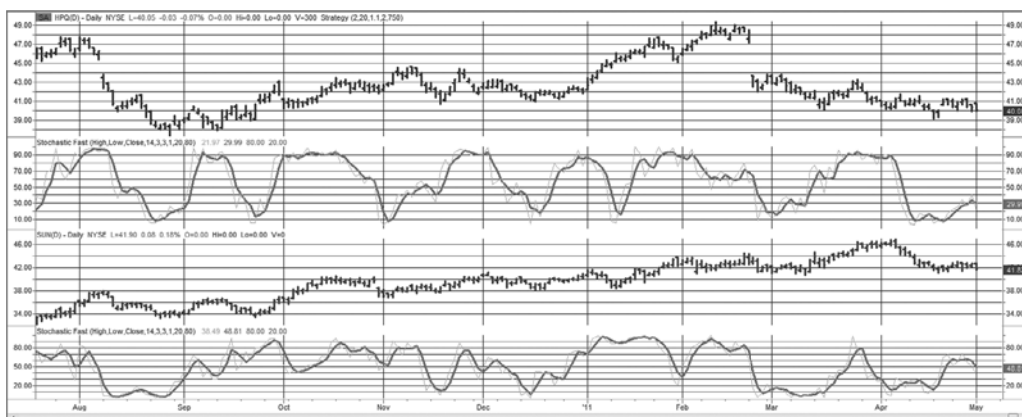
FIGURE 13.4 Hewlett-Packard (HPQ, top panel) and Sun Microsystems (SUN, third panel) with the 14-day stochastic indicator below each price series.

The stochastic indicator, or any other momentum indicator such as RSI or MACD, will show when the prices of two markets diverge. In Figure 13.4, Hewlett-Packard (HPQ) is in the top panel with its standard 14-day stochastic in the second panel, and Sun Microsystems (SUN) in the third panel with its stochastic at the bottom. The two companies are in a similar, but not identical, business. However, when the stochastics are compared, there are periods that are very similar and others that are not. There is a large drop in the price of HPQ in February 2011 and a much smaller drop in SUN, yet the stochastics both show a similar decline.

The stochastics can be used to systematically find trading opportunities. Because this trading approach is mean reverting, it is important that trades be held for only a few days. Anything longer will find that the longer-term direction of each price series will fight with the short-term corrections and make pairs trading profits more difficult.

Method 1: Difference between Two Momentum Indicators

Using the two stochastics, find the days on which the difference between the stochastics is greater than T, a threshold value. The maximum value of T is 100 and the minimum −100. A typical entry condition would be greater than 50 or less than −50. Exit when the difference in the stochastics returns to zero. Figure 13.5 shows the two stochastics together in panel 2 along with a third *stress indicator* discussed in Method 2.

Looking only at the two stochastics (the thinner lines in panel 2 of Figure 13.5), we can see periods where both stocks move together, the week of March 21, and when they are exactly opposite, such as the weeks of January 10 and March 7. When the HPQ stochastic is high and SUN low, we would sell HPQ and buy SUN, which was the case on January 10, while we would buy HPQ and sell SUN during the week of March 7.

Method 2: Stress Indicator

Using the difference between the stochastic indicators is one way to identify entry opportunities, but if those two values are combined into a single *stress indicator*, those entry
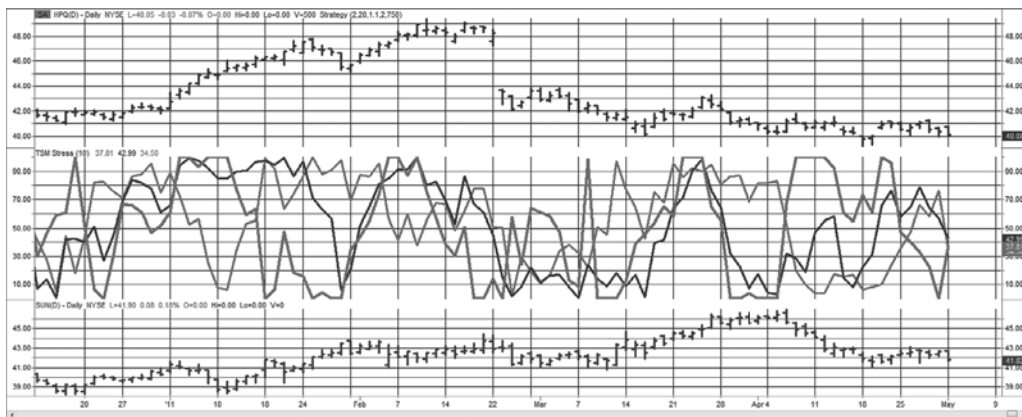
FIGURE 13.5 Hewlett-Packard (top panel) and Sun Microsystems (bottom panel) with the corresponding 14-day stochastics and the stress indicator in the center panel.

and exit points can be clearer. The stress indicator takes the difference between the two stochastics and uses that series as input into another stochastic calculation, which is the stress indicator. Note that only the *raw stochastic* is used.

$$Stochastic1_t = \frac{C1_t - \min_{i=t-n,\, t-1}(L1_i)}{\max_{i=t-n,\, t-1}(H1_i) - \min_{i=t-n,\, t-1}(L1_i)}$$

$$Stochastic2_t = \frac{C2_t - \min_{i=t-n,\, t-1}(L2_i)}{\max_{i=t-n,\, t-1}(H2_i) - \min_{i=t-n,\, t-1}(L2_i)}$$

The numerator is the distance between today's close and the low of the previous n days. The denominator is the maximum range of the previous n days. The stress indicator is

$$Stress_t = \frac{D_t - \min_{i=t-n,\, t-1}(D_i)}{\max_{i=t-n,\, t-1}(D_i) - \min_{i=t-n,\, t-1}(D_i)}$$

where $D_t = Stochastic1_t - Stochastic2_t$

The stress indicator can be seen in the center panel of Figure 13.5. Instead of looking for the difference of two stochastics exceeding a threshold, it is only necessary for the stress indicator to go about 95 (sell short leg 1 and buy leg 2) or below 5 (buy leg 1 and sell short leg 2).

A spreadsheet containing all of the calculations in the pairs section can be found on the Companion Website as *TSM pairs and stress calculations for HPQ-SUN*. This spreadsheet also contains the FX conversion rates to allow stocks or futures in different countries to be substituted.

Position Sizing

An important part of success in pairs trading is making the two legs equal in risk. If you were to trade one leg at an average price of $5 and another at an average of $25, you could expect the $25 stock to be much more volatile. Trading 100 shares of each stock would effectively make your returns entirely dependent on the success of the $25 leg.

The easiest way to equalize the risk is to calculate the position sizes based on the same investment, for example $1000. Find the average true range over the past 20 days and divide the $1000 by the dollar value of the average true range. Therefore, if the $5 stock had an average range of $0.25, we would buy or sell short 4,000 shares. If the $25 stock had an average range of $1, we would buy or sell short 1,000 shares. Using this process, both legs have an equal chance to contribute to the bottom line. When portfolio allocation and risk is discussed in Chapters 23 and 24, we will have a difficult time improving on this simple method.

Stock and Futures Relationships

When a listed company has a highly focused business dependent upon one or more physical commodities, it can be helpful to view both the price of the stock and the price of the commodity at the same time. As an example, Figure 13.6 shows the price of gold futures along with the share price of Barrick Gold (ABX), a mining company.

Although this chart would have been clearer if both gold and the stock price were expressed as percentages, we can show that the volatility of both markets are roughly the same. Gold prices span 1150 to 1550 and Barrick from 40 to 56. If we divide those ranges by the high price of each market we get 25.8% for gold and 28.5% for Barrick Gold. There is also great similarity in price patterns with the exception of the very end, where gold reaches new highs and Barrick drops on disappointing earnings. Once these prices
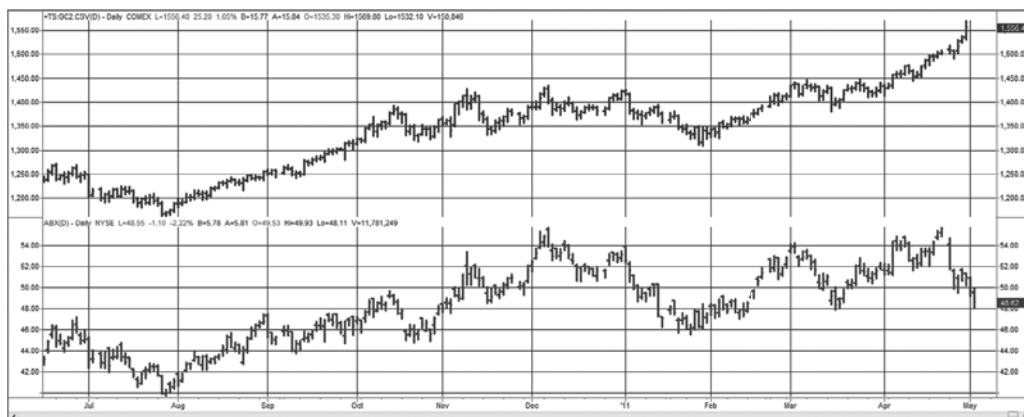
FIGURE 13.6 Gold futures (top) and Barrick Gold (ABX) from June 2010 through April 2011.

are converted to stochastics, the buy and sell signals are performed in exactly the same way as with only stocks. One advantage of trading gold futures is that you can buy or sell (short) with no restrictions and no added cost. If you choose to take only the signals where you bought Barrick and sold gold, you would not have to deal with short sales in the stock market. Of course, in exchange for convenience you would be giving up opportunity and diversification.

In the world of technical adaptation, it should be noted that pairs trading can be applied to options in a technique called *volatility dispersion trading.* Rather than taking advantage of differences in price, it looks for relative value differences in implied volatilities between an index and a basket of component stocks, targeting large differences. It more often involves short options positions.

Product Spreads

Product spreads take advantage of the relationship between the raw product and the results of processing that product. They are very actively traded. The *soybean crush* and the *energy crack* spreads are the most popular among futures traders and commercials, and have been used for decades. Product spreads usually involve three markets, which may be of different delivery months, depending on the processing time. The most common product spreads are:

Raw Product(s)	Resulting Primary Product(s)
Soybeans	Soybean meal and soybean oil
Crude oil	Gasoline and #2 heating oil
Feeder cattle and grain	Fat cattle
Feeder pigs and corn	Live hogs
Live hogs	Pork bellies

Exchanges recognize various spreads, executed in a specific proportion, as commercial application. They often allow reduced margins.

The Soybean Crush

Typically, the product spread is viewed from the perspective of the processor. A. E. Staley or Archer, Daniels, Midland (ADM) will buy soybeans, crush them into meal and oil, and then sell the products. You would expect the combined cost of the final products to be greater than the cost of the soybeans plus a margin of profit (the *gross processing margin*). Due to processing time, the crush relationship should be reflected in the current soybean price and the deferred product prices. When the gross processing margin is applied to futures prices, it is called the *Board Crush*, after the Chicago Board of Trade.

Processing one bushel of soybeans weighing 60 pounds yields 11 pounds of soybean oil and 44 pounds of 48% protein soybean meal, 3 pounds of hulls, and 1 pound of waste. Then

$$GPM = 0.022 \times Soybean\ meal + 11 \times Soybean\ oil - Soybeans$$

where meal is expressed in dollars/ton, oil in dollars/pound, and soybeans in dollars/bushel, the same units as the futures contracts. The exact ratio for trading would be 10:11:9 (soybeans:meal:oil); however, most traders use 1:1:1. Applying this to the current prices of November soybeans and December meal and oil on October 22, 2003, the GPM becomes

$$GPM = 0.022 \times 240.30 + 11 \times .2669 - 7.625 = 0.5975$$

Therefore, the GPM is nearly 60¢/bushel.

When the combined return on soybean meal and oil is below the cost of crushing the soybeans, processors may execute a *reverse crush* spread, in which they sell the soybeans and buy the products. Even though this appears to be a clever way of keeping prices inline, it is not done until the crushing margin is very negative. Processors cannot readily reduce their level of operation and lay off employees; a reverse crush means that they are buying products as well as producing them—a position of significantly increased risk.

The Crack Spread

Crude oil is refined primarily into gasoline and heating oil. As with the soybean crush, where the processor buys soybeans and produces meal and oil, the refiner expects a profit from the business of buying crude oil and selling gasoline and heating oil. A speculator can theoretically participate in this process by buying crude oil and selling the products in the right proportion and in the correct delivery months. For example, on May 5, 2011, crude oil for June delivery was $101.13 per barrel, heating oil $2.9338 per gallon for July delivery, and gasoline $3.0491 per gallon for July delivery. Based on futures contracts of 1,000 barrels of crude and 42,000 gallons of heating oil and gasoline, we could calculate each contract value:

	Price	Size	Value
Gasoline component	$3.0491	42,000	$128,062
Heating oil component	$2.9338	42,000	$123,220
Total products			$251,282
Crude oil component	$101.1300	1,000	$101,130

The other important fact needed to put this trade together is that it takes from four to six weeks to refine crude oil into its products; therefore, the prices used in the crack spread should always take the product prices quoted one month after the crude price. In the example above, June crude was paired with July heating oil and gasoline. Although there are many other byproducts of the refining process, the major components of fuel oil and gasoline are produced in a ratio of slightly more than 2 parts gasoline to one part fuel oil on average each year. This ratio then accounts for the 3-2-1 crack spread, where 3 contracts of crude oil are bought, and 2 contracts of gasoline and 1 contract of heating oil are sold. The exact historic ratios can be found in the *CRB Commodity Yearbook* under

Fuel Oil and Gas Production and Ratios. In the example above, we can complete the net transaction as:

	Trade	Price	Size	Value	Cost
Gasoline component	Sell 2	$3.0491	42,000	$128,062	$256,124
Heating oil component	Sell 1	$2.9338	42,000	$123,220	$123,220
Total products				$251,282	
Crude oil component	Buy 3	$101.1300	1,000	$101,130	$303,390
GPM (Net gain)					$75,954

Transaction costs will include execution slippage, commissions, and the cost of carrying the contracts until delivery (margin deposit and interest).

Reverse Crack

Similar to the reverse crush, when the delivered price of the products totals to less than the cost of crude oil, the refiner has no reason to process the crude oil, and can go to the market to buy the products and sell crude. This is called a *reverse crack*. In reality, processors can perform a reverse crack to some degree, but in the same way as the soybean processors, shutting down operations to accommodate a temporary market condition is not a good policy. The process of shutting down and restarting a large refinery is not simple, and labor problems make matters more complicated. When executing a reverse crack, the same delivery months are used for both crude and products.

Intercrop Spreads

A special case involving carrying charges is the *intercrop spread*, which can be highly volatile, even though there is an old crop inventory *carry-over* that ties the two seasonal markets together. Soybeans, for example, are harvested mainly in September and October. The August delivery is clearly the old crop, and November is the first new crop month; the September contract often reflects the shift from old to new and is avoided by most traders.

Normally, the old crop trades at a premium to the new crop. Carrying charges, accumulating since the previous winter, are part of the August price; export demand may cause shortages in the old crop, which move prices further above the cost of carry. Figure 13.7a shows the anticipated price pattern resulting from carrying charges during a normal year. The minimum storage commitment for the 3 months immediately following harvest is shown as a larger increase in price. Normal carry adds equal amounts to the price until the following September, where old and new crop mix. Finally, any carry-over (old crop inventory) must assume the price of the new crop, which is in greater supply.

The theory behind an intercrop spread is that prices must come together when the old crop merges with the new crop. But can it be a profitable trade? Examine the two possible events which affect this trade:

1. Problems involving development of the new crop making supply uncertain. This causes prices in the new crop to rise faster than the old crop (Figure 13.7b).
2. Export demand in the old crop, before the new crop is available, results in old crop prices rising faster than new crop prices (Figure 13.7c).

In Case 1, the spread between crops narrows and can only be traded as a spread that is *short the old crop, long the new crop*. This spread has limited potential since the November delivery cannot exceed the old crop by more than the normal carrying charge. At that point, processors will buy, store, and redeliver the old crop against the new crop. When prices in the new crop are nearing the old crop value, it makes sense to liquidate or even to reverse the spread. Once reversed, there is little risk that the spread can move adversely. A revised crop estimate that showed better yield than originally expected would cause the new crop to decline sharply and the spread to widen.
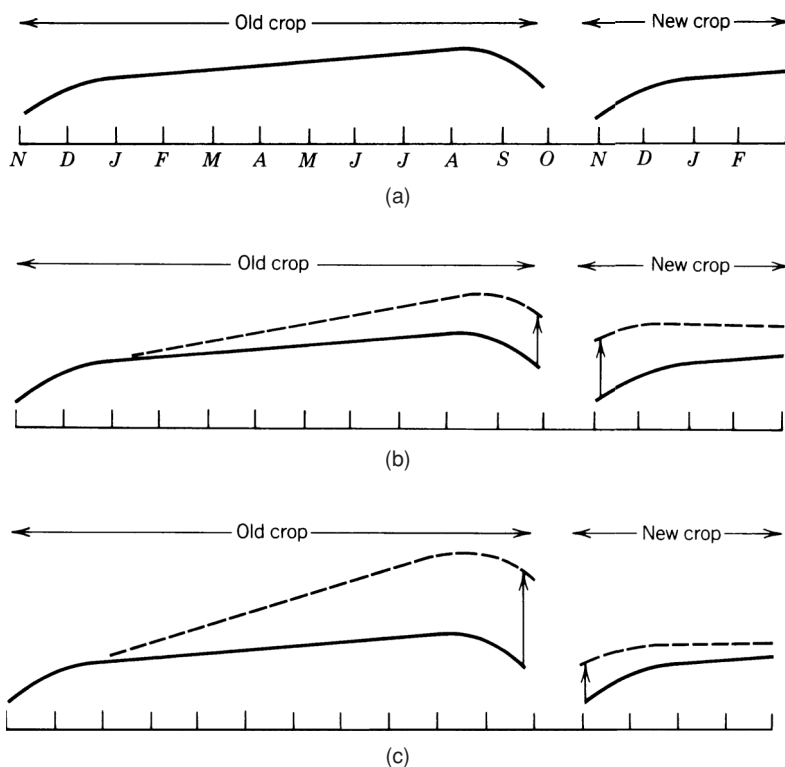
FIGURE 13.7 Intercrop spreads. (a) Normal carrying charge relationship. (b) New crop supply problems result in narrowing spreads. (c) Export affects the old crop more than new, resulting in a widening spread.

In Case 2, export demand in the old crop causes the intercrop spread to widen at first. A spread trader will typically wait until export commitments are complete, and then sell the old crop and buy the new. When old crop prices rise to a large premium over the new crop, many processors will reduce purchasing based on:

- Low or negative profit margins at the current price levels.
- Use of reserves or inventory to carry processing through until the lower new crop prices are available.

It is remarkable how demand is inversely related to price even when it is considered *inelastic*. When a delay in purchasing or processing will result in greater profits for the commercial, that delay is somehow achieved.

Butterfly Spreads

A *butterfly spread*, normally referred to as just a *butterfly*, is a low-risk technique for capturing short-term price distortions between delivery months in the same futures markets (the *term structure*). It is best in highly volatile markets, where the concentration of trading in one or two delivery months causes those contract prices to move away from the normal delivery month relationship. These distortions often occur in the week before the nearby futures contract expires, when large positions holders, such as fund managers, roll from the nearby month to the next delivery. Opportunities most often exist during the trading day; however, setting a butterfly is the same process whether looking at intraday or closing prices.

For example, in April a combination of poor planting, declines in the U.S. dollar, and new export agreements changes the soybean prices as shown in Figure 13.8. Increased demand results in sharply higher July futures prices with a tapering-off of the effect in the more deferred contracts. The normal carrying charge relationship is shown as a straight line in the old crop, beginning again in the new crop. A butterfly entered in the old crop would mean selling two contracts of July soybeans and buying one contract each of May and August. This is the same as executing two spreads: long May, short July, and short July, long August.
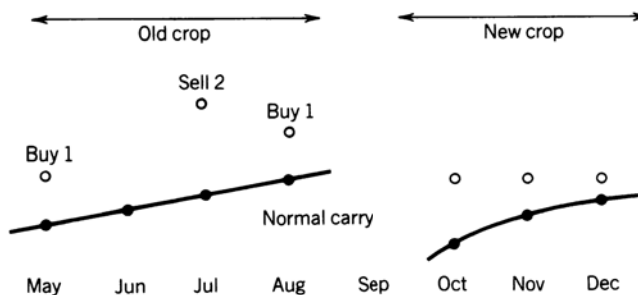
FIGURE 13.8 Delivery month distortions in the old crop making a butterfly spread possible.
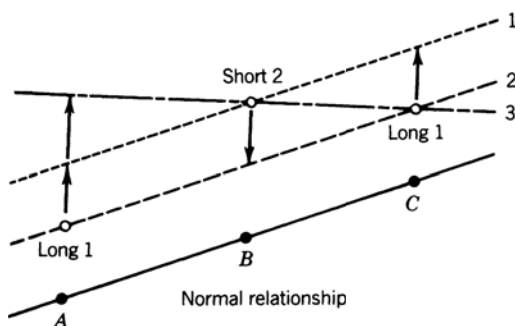
FIGURE 13.9 Corrections to interdelivery patterns.

Each spread in the butterfly has a good chance of being profitable; the combination of the two is exceptionally good. The July contract cannot remain out of line with the deliveries on both sides because a trader could take delivery of the May contract and redeliver it in July at a profit exceeding the cost of carry. Under normal circumstances, commercial users of soybeans will defer their purchases to later months, depleting their reserves, to avoid paying a short-term premium and causing July prices to drop.

Another reason for distortions in the *term structure* of the contract is concentrated orders by large hedge funds, typically macrotrend trading programs. Those funds tend to trade in the most liquid months and many do not make an effort to spread their orders across more than one delivery. Instead, they let the market makers, or the floor brokers, execute their orders by spreading against other delivery months. If a large buy order is placed in the May soybean contract, the broker will sell May and buy July, and if possible also buy March and sell May, creating a March-May-July butterfly. Brokers will rarely hold an outright long or short position.

The butterfly spread guarantees that any adjustment in the three contracts back to a normal relationship will prove profitable. If the price of both May and August rise to be in-line with July or if May rises to form an inverted relationship, the spread will be profitable (Figure 13.9).

The problem with such an ideal spread is the short window of opportunity and difficulty executing the orders at the right moment. Because profits are nearly riskless, opportunity is small. The beneficiaries of these trades are usually the floor traders who can act quickly. Once the position is entered, liquidation can be easily accomplished.

A common problem for those attempting butterfly trading from screen prices is that lack of trading in the deferred months may appear to be a distortion. Prices may be bid or asked at the correct level, but not traded. If the bid-asked prices do not show on your screen, then the last traded price that is out of line may be a false opportunity.

The Yield Curve

There is a wide range of strategies involving the yield curve, or term structure, of various futures markets. At the most sophisticated level, the large investment houses, hedge
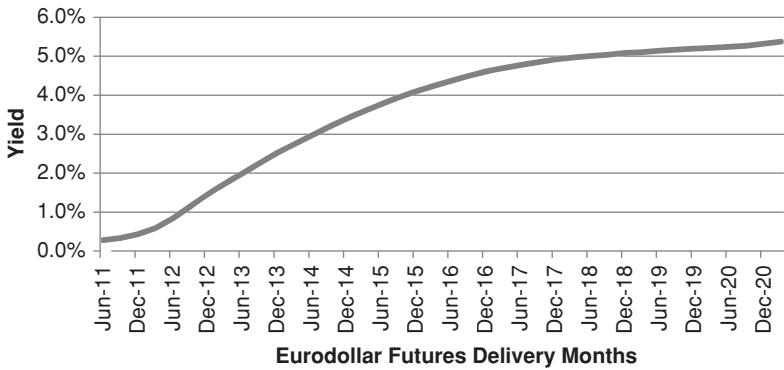
FIGURE 13.10 Eurodollar term structure of futures prices.

funds, and commercials make a business of arbitraging individual Eurodollars futures prices against strips of the same maturity as well as against comparable bond maturities, normally the 2-, 5-, and 10-year Treasury notes. To be successful, these traders own seats on the exchanges so that their transaction costs are near zero.

A *Eurodollar strip* is the partial returns of each Eurodollar futures contract combined beginning with the cash Libor yield and then compounded with the partial returns of each futures contract. This can be expressed as[8]

$$TV_T = \left(1 + L_0 \frac{D_0}{360}\right) \times \left(1 + F_1 \frac{D_2}{360}\right) \times \cdots \times \left(1 + F_n \frac{D_n}{360}\right)$$

Where TV_T = terminal value of \$1 invested for T years

       L_0 = spot Libor rate

       F_i = futures rate for the 1st deferred contract (100 – futures price)/100

       D_i = actual days in each period

Figure 13.10 shows the term structure of Eurodollar futures on April 5, 2011. As an example of an arbitrage possibility, consider the first two futures contracts. If the cash Libor rate is 0.21 and there are 35 days until the nearby Eurodollar futures contract expires, the rate of the first futures is 33 basis points (0.33%), and the rate of the second futures is 0.43%, then the futures price of the second can be calculated as

$$TV_{F2} = \left(1 + 0.0021 \frac{35}{360}\right)\left(1 + 0.0033 \frac{91}{360}\right)\left(1 + 0.0043 \frac{91}{360}\right)$$

$$= 1.002139, \text{ or } 0.2139\%$$

---

[8] Galen Burghardt, *The Eurodollar Futures and Options Handbook* (New York: McGraw-Hill, 2003).

The futures price would then be $100 - 0.2139 = 99.78608$. The actual futures price is 99.57, giving a differential of about 21 basis points. For an arbitrage to be successful, traders look for a minimum of 25 to 30 points. Note that 91 days were used for each futures market, although the actual days may vary slightly and should be used.

The Eurodollar strips, as well as the Eurodollar futures price, for the 5-year forward delivery should also be equal to the yield on the 5-year note or any note that is due to mature in the same time. Some threshold profit level is needed to cover the cost of holding the Eurodollar contracts, the note or bond contracts, or to execute the cash market transactions. To make this easier, the Chicago Mercantile Exchange (CME) has initiated a contract called a *Eurodollar bundle*, which trades from 1 to 10 years, essentially creating strips of those maturities from Eurodollar futures. This removes the need to execute all delivery months when creating a strip. It is implicit in the new contract.

Trading the Yield Curve

Many traders see shifts in the yield curve, steepening or flattening, as a trending process. As the central bank lowers rates in the near-term, the yield in the short end declines more than the long end, and the curve steepens. Trend traders try to exploit this by creating spreads between the various interest rate maturities, Eurodollars, 2-year notes, 5-year notes, 10-year notes, and 30-year bonds. This can be done simply by creating a new series of, for example, 10-year notes minus 5-year notes, then trading that series by

> *Buying* 10-year notes and selling 5-year notes when the yield curve is steepening, and
> *Selling* 10-year notes and buying 5-year notes when the yield curve is flattening.

The shorter maturities will tend to be less volatile than the longer end, therefore it is necessary to size your position to equalize the volatility; otherwise, your profits and losses will always reflect the performance of the longer maturity. To do this, we use the same method as discussed in pairs trading, take an equal investment amount for each market and divide by the dollar-value of the volatility, using the *average true range* of the past 20 days. The more volatile market will have fewer contracts.

It is difficult to include the Eurodollars in this type of trade because prices are not quoted in the same terms as the other interest rate markets. You would want to convert all prices to yields.

The TED Spread

The difference between 3-month U.S. Treasury bills and 3-month Eurodollars is traded as the *TED spread. Eurodollars* are U.S. dollars on deposit in banks not in the United States and held by both non-U.S. banks and branches of U.S. financial institutions. While originally only European banks were included, Eurodollars can now be in any major bank in the world.

The TED spread is considered an indicator of credit risk and represents the shifting public confidence between the U.S. government and the highest-quality non-U.S. domiciled banks. During times of crisis, money flows back to the United States as a safe haven.

The TED spread is traded in futures of the same delivery month. Typically Eurodollar rates are higher; therefore, if March T-bills are at 97.50 and March Eurodollars are quoted at 97.00 (discounted rates of 2.5% and 3.0%, respectively) the TED spread is 0.50, quoted simply as the whole number 50. The TED spread rises when there is greater economic uncertainty. Banks outside the United States offer higher rates to prevent money from flowing back to the U.S. Treasury. The U.S. T-bills will always trade higher (a lower yield), offering lower rates but more security.

Implied Interest Rates

After the subprime crisis in the fall of 2008, many investors were concerned about the counterparty risk even with large banks. After all, if Lehman Brothers could fail, why not any bank or investment house? When you buy a Certificate of Deposit (CD) issued by a bank, you are lending them money based entirely on the credit-worthiness of that bank. Sometimes that loan is insured by a third party, but not often, and the cost of insurance reduces the return of the CD. In a worst-case scenario, the insurance is written by AIG, a company that was itself on the verge of collapse. Instead, many investors looked to the futures markets, which are guaranteed by the clearing house. While there has never been a default in futures, it is always possible that the level of insurance provided by the clearing house will not be enough in a crisis. An ongoing investigation into missing funds at MF Global may come to rest at the CME Clearing Corporation. However, it is considered better than the alternatives.

When a futures market trades in contango, where each month is higher than the previous month, we can assume that most of that increase is due to carrying charges. This does not apply to agricultural markets, but the main focus is on extracting the dividend yield from stocks or the implied yield from precious metals. While this may be less than the current CD rate, the investment is smaller (margin rather than full cash) and seen as more secure.

A *pure carry* market is one in which the forward price is entirely comprised of the costs of holding the physical product until a predetermined delivery date. As previously mentioned, these costs are those of storage, insurance, and the loan rate. As an example, consider gold, which is a classic pure carry market.

On April 5, 2011, the price of the December 2011 contract was 1484.9. If we want to know much interest is gained by buying the December contract, and simultaneously selling the December 2012 contract at 1496.9, we can apply the normal calculation for annualized rate of return (see Chapter 2), using the actual elapsed days (in this case 362), and get 80.37 basis points. These returns can be derived from the numbers in Table 13.2. Eurodollar futures show that the bank returns for the same period would be 43 basis points. In addition, to extract the 80 basis points, you need only invest the required margin plus some amount of reserves rather than the entire face value of a CD. Figure 13.11 shows the gold futures prices and the implied yield based on the first May 2011 contract and the differences to each of the forward deliveries. A simple program to calculate the implied interest rate, *TSMimpliedinterest*, is available on the Companion Website.

TABLE 13.2   Gold Prices and Delivery Months, Implied Yield, and Total Days

Month	Price	Implied Yield	Days
May-11	1480.9		
Jun-11	1481.4	0.004059	30
Aug-11	1482.6	0.004549	91
Oct-11	1483.8	0.004675	151
Dec-11	1484.9	0.005350	182
Feb-12	1486.1	0.005228	242
Apr-12	1487.6	0.005396	302
Jun-12	1489.1	0.005506	362
Dec-12	1496.9	0.007137	544
Jun-13	1509.8	0.009630	726
Dec-13	1529.1	0.012780	908
Jun-14	1551.6	0.015522	1090
Dec-14	1577.9	0.018118	1272

THE CARRY TRADE

A strategy that gets blamed for many of the large moves in currencies is the *carry trade*. The idea behind this multimarket approach is simply that money flows to the countries that have the highest interest rate yield net of inflation, that is, *real returns*. For example, if the interest rate in the United States is 4% and the rate in Japan is 1%, as it was for many years, then we would borrow Japanese yen, convert to U.S. dollars, and buy U.S. Treasuries for a gain of 3% per year. When we are done, we would return the amount of yen originally borrowed. The risk in the trade is that the U.S. dollar will weaken against the yen (USDJPY) causing an offsetting loss in the currency transaction that could easily be larger than the expected gain in interest income.
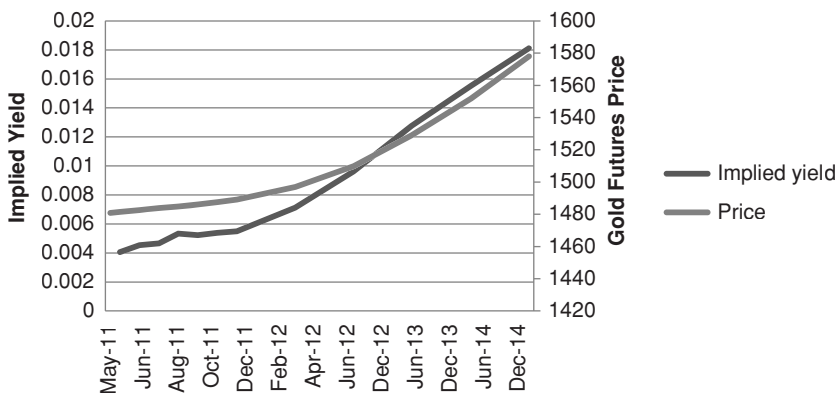
FIGURE 13.11   Implied yield from gold futures, April 5, 2011.

Although it is thought to be extracting the carry, the cost of holding an asset for some time period, it is a misnomer. Traders do not own cash yen and exchange that for cash dollars, and buy cash bonds in order to extract the carry. Instead they understand that the interest rate differential between the United States and Japanese 3-month returns will translate into *interest rate parity* (discussed earlier in this chapter) and the forward price of the dollar versus the yen will be higher if the dollar interest rate is higher than the yen interest rate. Therefore, they focus entirely on the futures or forward markets, where they can invest with exceptionally high leverage. Then, if the interest on a dollar investment is greater than the interest on a yen investment, they buy the dollar-yen, USDJPY. But, the forward price, by definition, discounts the forward interest rate differential, so how can you capture the carry if it's already discounted in the futures price?

The answer is called *forward bias*. Because the trade is executed in the forward market, it cannot extract the carry. Instead, it relies on the time-tested phenomenon that money flows to the country with the highest real rate of return. There is high confidence level that the U.S. dollar, yielding 4%, will strengthen over the yen yielding 1%, or that Brazil at 8% will strengthen over the dollar at 4%, providing Brazilian inflation is less than 4% and that the economy is otherwise stable.

How significant is the risk that the currency will move against you? That the yen will strengthen against the dollar even when U.S. rates are much higher than Japanese rates? Sometimes the risk is substantial; therefore, a diversified portfolio is needed to offset the times when some of the trades go the wrong way.

The Basic Carry Trade

Most traders would like to implement the carry trade using the major industrialized countries, G7. Those are Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States. Of those, there are now only 5 currencies. To be even more practical, it would be easiest to implement the trade using the actively traded currency futures markets, the euro (EURUSD), British pound (USDGBP), yen (USDJPY), Swiss franc (USDCHF), Australian dollar (AUDUSD), Canadian dollar (CADUSD), and perhaps the Mexican peso (USDMXN). Using only exchange-traded markets reduces the chance of counterparty risk. There are many other currencies listed on the CME that are illiquid, and they allow traders to *EFP* (*exchange for physicals*), that is, they can execute the currency trade in the cash market and exchange it for a futures position. At the beginning of each day the bank FX desk will post the number of basis points charged or credited in order to move the cash position to a futures position. It is simply the interest rate parity calculation, plus a transaction charge.

Now consider the difference in the spot FX rate and the nearby futures price for these crossrates. In Table 13.3, the spot and futures prices plus the corresponding cash 3-month interest rates are shown for February 27 and March 31, 2009. Cash rates can be found on Bloomberg, while spot and futures prices for crossrates can be found on CSI, Bloomberg, and other data services. Because the CME quotes some of these rates differently from the FX market, and the FX market is itself inconsistent in the

TABLE 13.3 Key Values of Crossrates and Yields, March 31, 2009

Country	Crossrate	27-Feb-09		31-Mar-09		27-Feb-09	Change in	
		Rate	Futures	Rate	Futures	Differentials	Rates	Futures
USA	USA	1.2644		1.1919			−0.0725	
EC	EURUSD	1.8250	1.2656	1.5100	1.3242	0.5606	−0.3150	4.63%
Switzerland	USDCHF	0.4950	0.8661	0.4033	1.1278	−0.7694	−0.0917	−2.32%
Canada	USDCAD	1.3500	0.7826	1.0000	1.2686	0.0856	−0.3500	−0.72%
Australia	AUDUSD	3.1633	0.5698	3.1383	0.6224	1.8989	−0.0250	9.23%
UK	GBPUSD	2.0494	1.4270	1.6488	1.4288	0.7850	−0.4006	0.13%
Mexico	USDMXN	7.5000	0.0588	6.5000	0.0639	6.2356	−1.0000	8.59%
Japan	USDJPY	0.6375	1.0319	0.6031	0.9821	−0.6269	−0.0344	−4.82%

way it quotes some of the rates, it is necessary to show all rates as "per USD" (how many USD are needed to get 1 of the other currency units), the values have been converted into a standard notation, shown in columns 5 and 6. The *interest rate differentials* are the difference between the 3-month rates of other countries compared to the U.S. rate. The last column shows the percentage change in the futures price from February to March.

Positions are taken in the carry trade to reflect the direction of the interest rate differentials, most often at the end of each month. Based on February 27 shown in Table 13.3, those would be

> *Long* the euro against the U.S. dollar
> *Short* the Swiss franc against the U.S. dollar
> *Long* the Canadian dollar against the U.S. dollar
> *Long* the Australian dollar against the U.S. dollar
> *Long* the British pound against the U.S. dollar
> *Long* the Mexican peso against the U.S. dollar
> *Short* the Japanese yen against the U.S. dollar

Each month new interest rates are recorded, differentials recalculated, and positions changed and rebalanced according to their volatility.

Although this month was not intentionally selected, 6 of the 7 crossrates satisfy the expectations of the carry trade, as shown in Figure 13.12. As the difference between the interest rate of one country gets larger compared to the U.S. rate (e.g., Mexico, the largest), the expected change in the currency also gets larger (although with considerable variability). The value of the Mexican peso increased from about 17 per 1 USD to 15 per 1 USD. Only the Canadian dollar, which had a slightly higher interest rate, lost value to the U.S. dollar, but then the rates were very close.

February 2009 was still part of the window affected by the subprime crisis, and both rates and futures prices were very volatile. Although it did not happen during this
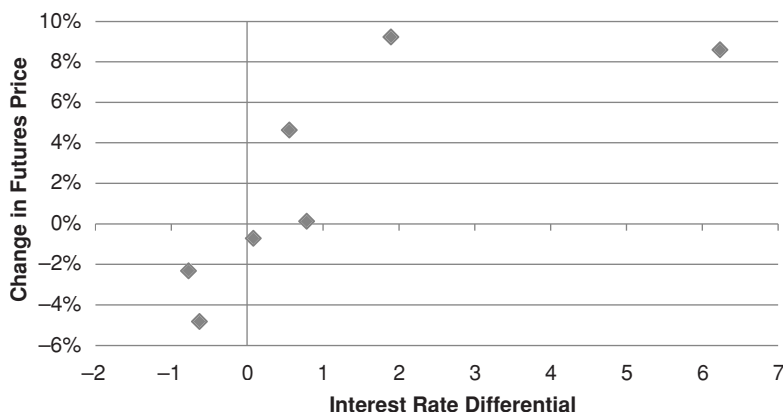
FIGURE 13.12 Interest rate differential versus change in futures prices during March 2009 for seven crossrates.

example, futures prices can move fast and further than most of the differentials (which are less than 1%), offsetting the perceived gain from the carry. In order to control the risk, most traders will volatility-adjust the position size of each currency position, improving the effect of diversification.

The biggest problem in the carry trade that uses the developed countries, which are the most liquid and traded on the CME, is that they have the smallest interest rate differentials and the most correlated currency price movements. When the United States releases a poor trade balance number, all currencies strengthen against the dollar. Adding the Mexican peso to the mix greatly improved returns because the differential was far larger than any of the other countries. However, as emerging markets are added, the risk of a large unfavorable currency move increases, as does the volatility of returns.

Another problem to be resolved is the concentration of positions in the *funding currency*. In Table 13.3 the U.S. rate was 1.26, the third-lowest, with Japan and Switzerland typically offering the lowest rates. The funding currency is the one that is shorted (or borrowed), so that for the choice of contracts on the CME, the carry trade would have shorted the U.S. dollar in five of the seven. That makes success highly dependent on the direction of the U.S. dollar, specifically that it continues to decline relative to the other currencies. If we were trading in the cash market, we could have sold the yen or Swiss against all the other currencies, but then there would be high concentration in those markets. And, while it's possible to create a mix of the major crossrates, there is still likely to be concentration in the funding currencies and more difficulty executing less active combinations.

The theory of forward bias continues to work, but with intervals of high volatility and drawdown. The performance record of major institutions, such as Deutsche Bank and Barclays, are available online.

CHANGING SPREAD RELATIONSHIPS

Spread relationships can change over time due to structural changes in both the supply and demand, as well as changes in consumer habits. Consider the following:

1. Seasonality for crops, orange juice, and many other products has changed as countries in both hemispheres increase production and join world export markets. Grain and orange juice from South America and edible oils from the Far East have permanently altered the seasonality of supply. This competition puts pressure on carrying charges.

2. Large agricultural conglomerates are able to use futures markets to hedge their production or purchases, shifting the need to buy or sell away from typical harvest patterns. An increase in the storage facilities also decreases the need to sell at harvest.

3. Traditional ratios between gold, silver, and platinum are changed by new demand in the industrial use of those metals in the electronics and automotive industries. The rise of the middle class in China and India has added substantially to demand, as well as the metals' perceived usefulness as a hedge against inflation and changes in currency values

4. Consumer tastes change, causing a shift away from red meat for some period, then back to red meat during another. Lean hogs have replaced the traditional hog futures contract. Health concerns about the use of nitrates in bacon will change processing or pricing; consequently, the ratio between hogs and pork bellies changes.

5. New products, such as high-fructose corn sweeteners, have reduced demand for sugar. The use of corn for ethanol has further increased demand.

Very little remains the same over time. Even though there may be anticipation of change, unexpected events relating to health or interruption of supply can produce a price shock which will affect many spreads.

In the stock market, the relationship between stocks in the same industrial group is affected by trading in index futures. If one stock belongs to the small caps and another to the large caps, their prices may be forced apart by buying in the Russell 2000 and selling in the S&P 500, and the subsequent arbitrage—moves that do not consider the individual stock relationships or fundamentals.
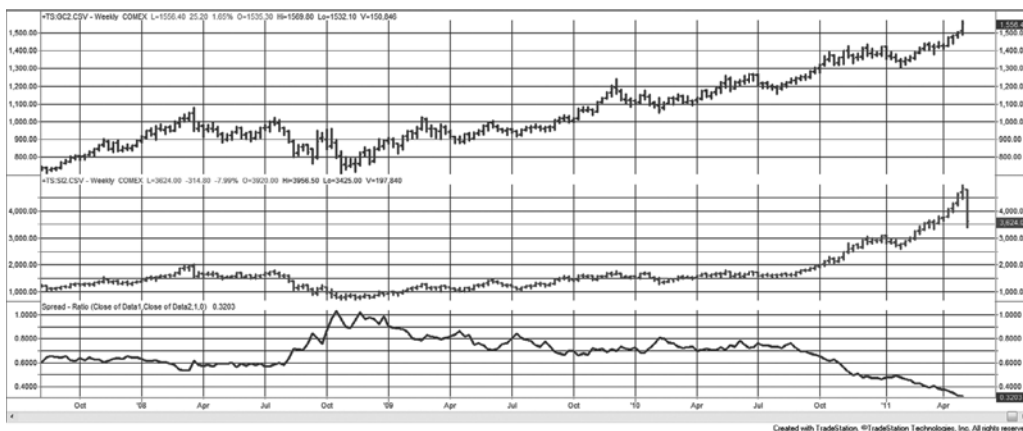
Gold/Silver and Platinum/Gold Ratios

A relationship that has always been followed with keen interest is the gold/silver ratio, traditionally considered normal at 33:1. Both markets have been considered a safe haven in times of economic instability. It serves as a good example of the variability of many perceived spreads. When gold was $35 per ounce and silver about $1 per ounce, the relationship was never very stable. Between 1930 and 1945, silver prices dropped to about $0.50 per ounce; from 1951 to 1962, they remained just under $1 per ounce and

afterwards began their accelerating move to $38 per ounce in January 1980. Between 1944, when the International Monetary Fund (IMF) was formed, and 1971, the price of gold was fixed at $35 per ounce although in the few years before the cancellation of the agreement, the price varied above the designated value. From the late 1960s until the silver crisis of 1980, the gold/silver ratio ranged from about 20:1 to 45:1 in a slow cycle. During the crisis attributed to the Hunt brothers (September 1979 to September 1980), the ratio fluctuated around 33:1 in a highly volatile pattern moving only slightly above and below its historic pattern. Following the silver crisis, the value of silver declined faster than gold reaching a ratio of 56:1 in June 1982, with gold at $314 and silver at $5.57. A year later, the ratio touched 33:1 and rapidly widened to 66:1 in June 1986.

Since late 2008, there has been a substantial increase in the price of gold due to fears of economic instability and high inflation. During that time, the price of silver lagged behind, with the ratio reaching over 100:1 during the crisis, then falling back to 70:1 (see Figure 13.13). As gold crossed over $1200/ounce silver suddenly rallied, bringing the ratio sharply down to the original 33:1, only to see a sell-off of nearly a third of its price. Although profits could have been made during the stable periods, such as 2010 through mid-2011, the stability of the ratio is all in the perception of the general public when they choose gold or silver as protection. Recently, gold has been proved the metal of choice.

The platinum/gold ratio (Figure 13.14) serves the same purpose as the gold/silver ratio, although gold and platinum seem to have found their way into commercial use and have had a more constant relationship than silver and gold. Structural shifts in the ratio are likely from time to time because of changing industrial uses. For now, gold is used as an electronic conductor and platinum has been in demand for automobile catalytic converters. There is no fundamental reason for these prices to remain in the same ratio going forward.

FIGURE 13.13 The weekly gold/silver ratio, using futures, was once normal at 33:1, but has gotten as high as 100:1 before falling back to 33:1.
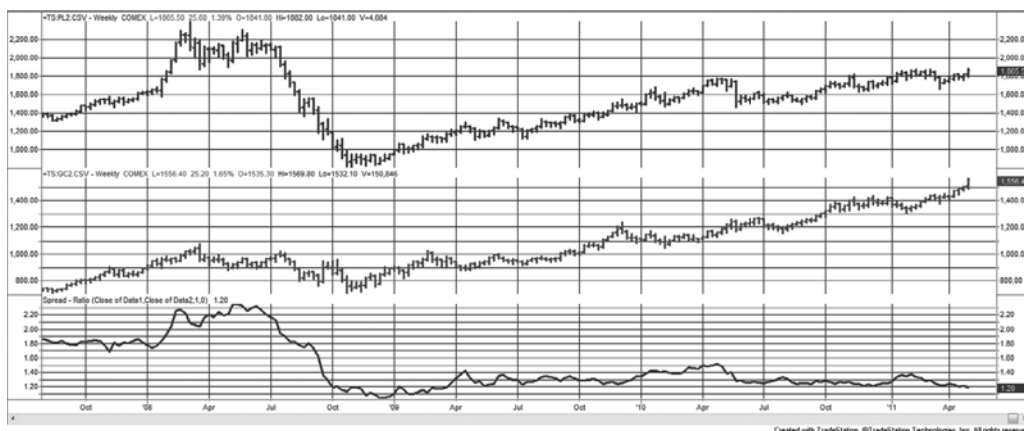
FIGURE 13.14   The platinum/gold ratio, shown here using futures, has been reasonably stable.

INTERMARKET SPREADS

There are many dependent relationships between different futures markets that are not classified as arbitrage. Because the bases for these intermarket relationships are fundamentally sound, spreads are an excellent way to profit from short-term divergence. The most comprehensive study of these relationships can be found John Murphy's *Intermarket Technical Analysis*,[9] in which he studies the leads and lags between markets and examines the interaction in their relationships. More recently, Markos Katsanos' *Intermarket Trading Strategies*,[10] offers a detailed study of these relationships.

Many readers will be aware of the increased correlation between many markets during the past 10 years. A combination of globalization, arbitrage, and the movement of money seeking better returns, or running from risk, has caused markets that had only a casual relationship to each other to often move together. Investments in large commodity funds (such as the Goldman Sachs Commodity Index) now contribute to sympathetic movements in grains, energy markets, and metals simply because of the flow of money.

The most elusive of the intermarket spreads is the relationship between bonds and the stock market, represented by the S&P 500. The Federal Reserve manages economic growth and controls inflation using interest rates; therefore, it is natural to expect the stock market to react to these orchestrated changes. As interest rates decline, the S&P should rise. In economics, the reactive phenomenon is described as a *J-curve* because there is a delay between the lowering of interest rates and the response by the stock market. The lower part of the *J* pattern indicates that prices continue lower at first but at a slower rate before eventually responding to the stimulus. In most cases, such as the

---

[9] John J. Murphy, *Intermarket Technical Analysis* (New York: John Wiley & Sons, 1991).

[10] Markos Katsanos, *Intermarket Trading Strategies* (Hoboken, NJ: John Wiley & Sons, 2008).
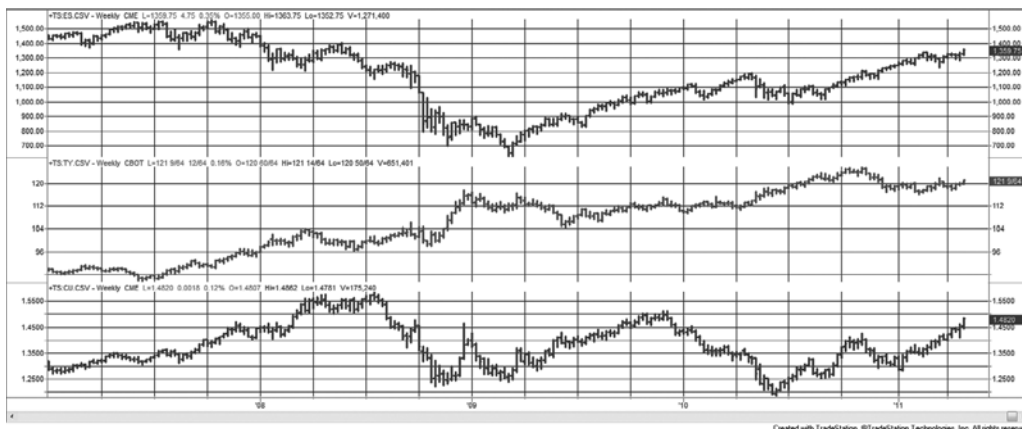
FIGURE 13.15 10-year notes (top panel) cause a delayed response in the stock market (center). The strengthening of the euro (bottom panel) will improve the U.S. trade balance.

recovery that began in 2002, the Fed needs to lower rates a number of times before a sustained positive reaction can be seen in the stock market.

The *J* pattern can be seen in Figure 13.15. The price of 10-year Treasury notes rises slowly during 2007 while the S&P declines. Interest rates continue sideways until September 2008 when the subprime crisis causes the Fed to drastically lower rates throughout the last quarter. After the stock market makes one more move down, it starts a long climb up as rates also drift lower (prices higher). The euro is more difficult to interpret. A higher euro indicates an economy that is seen as stronger than in the U.S. but at the same time helps U.S. exports. This also occurs with a delay; therefore, from the beginning of Figure 13.15 through mid-2008 we see the euro rising due to a poor U.S. economy reflected in declining S&P prices. After the 2008 crisis, the interpretation of which economy is stronger has fluctuated back and forth, but persistent lower interest rates in the United States and failure to deal with its debt problems have favored the euro. All else being equal, such as inflation and government stability, the values of currencies respond primarily to interest rates. Investors move funds to capture the highest real rate of return, the interest rate net of inflation.

There are many other intermarket spreads that should be watched:

- All of the U.S. index markets, the DJIA, S&P, Russell, and NASDAQ
- The European index markets, the EuroStoxx, DAX, and CAC.
- The London Metals Exchange (LME) nonferrous metals, copper, tin, lead, aluminum, zinc, and nickel. All are used in housing.
- Gold and crude oil, one a major component of inflation and the other a hedge against inflation.
- Commodities denominated in gold or other currencies, giving us a view of what others pay for a product and creating a new market at the same time.

Increased Spread Risk

While any spread trading appears to limit risk because it reduces the net price move when the markets being spread are clearly related, it is not usually true. *Cross-margining* recognizes that buying one market and selling another related product has intrinsically lower risk than simply buying or selling short in a single market. The brokerage firm that you use will know that a long in S&P futures and a short in the Dow futures is a reduced-risk position; therefore, they will give you favorable margin consideration. Margin on these two trades will usually be reduced to the larger margin requirement of the two trades had they been taken as separate positions. Margin represents a measure of risk. When entering an intradelivery spread, for example, long March and short June 10-year Treasury notes, the margin can be remarkably small because the risk is perceived to be small. Lower margin equates to higher leverage and increases the risk for most traders who measure the returns on the trade as a percentage of the margin requirement. In some cases, the spread risk will be greater than the risk of an outright long or short position. When prices diverge, the leverage can work against you.

Spreading two markets of different denominated currencies, such as the S&P and EuroStoxx will add the uncertainty of changing FX rates. Whenever you add a third uncertainty, the risk increases.

Creating New Markets to Trade

Until a few years ago, Russia bought all of its oil in gold. Oil is typically quoted in U.S. dollars. When a country imports any product, they use their own currency, or are required to pay in the currency of the country selling the product or in a specified third currency. Sometimes the decision to buy, that is, the timing of when to buy, depends on how the price is seen denominated in the currency of the buyer. If we divide the price of oil by gold or by the euro, we see a different picture of price movement. It effectively becomes another market to which we can apply trend following or other trading methods and get returns that add diversification to a portfolio.

As an example, we will look at oil/gold. In Figure 13.16, crude oil weekly futures are shown at the top, gold weekly futures in the middle, and the ratio of oil to gold at the bottom. A 40-week moving average appears in each panel. The important point is that the moving average based on the ratio turns down much sooner than either of the individual markets and captures a significant profit in the ratio. The creation of a ratio has essentially created a new market in which trend trading signals occur at different times from either of the two underlying markets.

Trading the ratio is different from trading the difference between two markets, and the ratio may have problems if you use the oldest part of a back-adjusted futures series because those prices can approach zero or even go negative. If we look at only the data from 2000 forward, the results are safe. At the far right of Figure 13.16, the beginning of 2008, the price of crude is $135, and gold is $950, giving a ratio of 0.19. We will create a simple trading strategy that buys crude and sells gold when the ratio is rising,
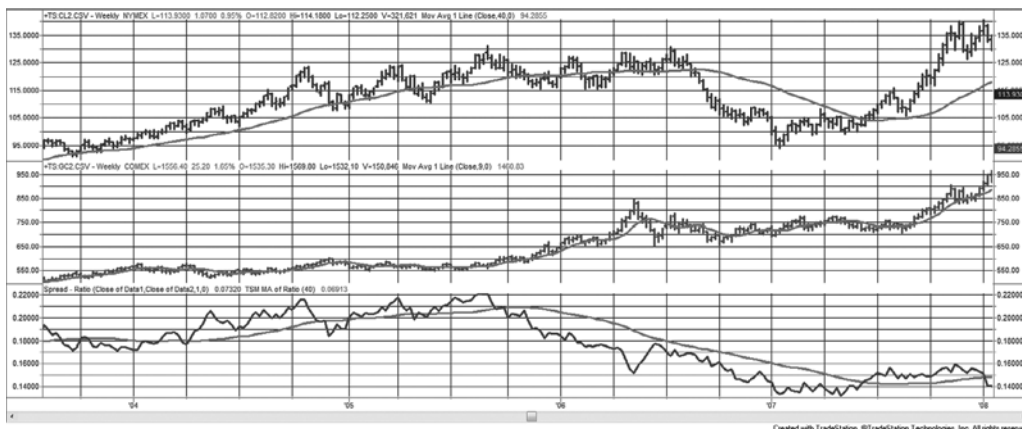
FIGURE 13.16 Crude oil (top) denominated in gold (center, both weekly futures) gives the ratio at the bottom. A 40-week moving average is shown in each panel.

and sells crude and buys gold when the ratio is falling. Each time the strategy enters a trade, its volatility adjusts both legs in the same way discussed in the "Pairs Trading" section of this chapter. Alternatively, we could find the number of contracts of gold that exactly hedges the value of the crude oil contract, which is what would happen if we were buying physical crude oil. Given crude at $135 and gold at $950, we would be buying or selling 1.42 contracts of gold for every 1 contract of crude, or a ratio of about 3:2. However, for this example, the spread legs will be volatility-adjusted in order to minimize the risk.

A comparison of the performance of this combination can be seen in Figure 13.17. Part a shows the results of signals based on the ratio, with the cumulative profits of each leg shown along with the net results. Part b shows the results of trading crude and gold individually, also with the position sizes volatility-adjusted. Clearly, the results of the ratio are lower than the combined individual results. However, the volatility is also lower. If we adjust the volatility of the ratio method up by a factor 2.64, the results are then on an equal footing. Still, the net returns of the individual markets are nearly twice that of the ratio.

These results are simply one test with a 40-week average; therefore, they may not be representative of the comparative performance under a broader set of tests. The purpose of the ratio was to create a new market with results that would add diversification to a portfolio based on a trend-following strategy. To see if this was accomplished, the correlation of the weekly profit or loss changes was found to be 0.42, reasonably low. Remember that, if the two methods are both trend following and they are both profitable, then they must be in the same positions during profitable periods, so a correlation below 0.30 is unlikely.

A clearer test of whether the ratio method offers diversification is to compare the actual entry prices of the two methods. As you can see in Figure 13.18, the ratio method

Performance of Ratio


(a)

Performance of Individual Series
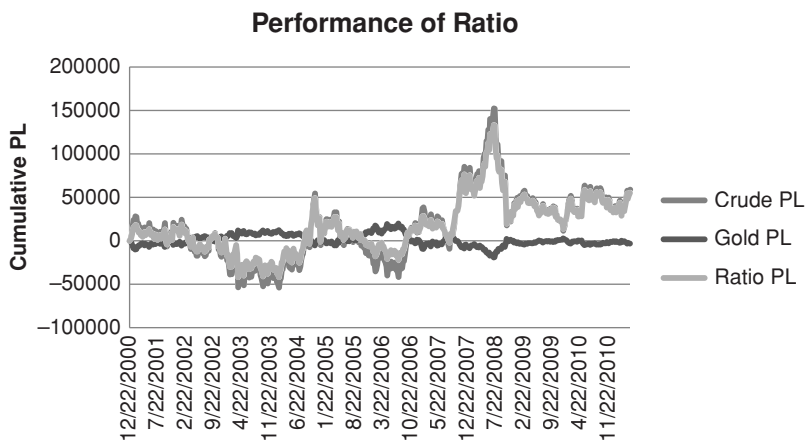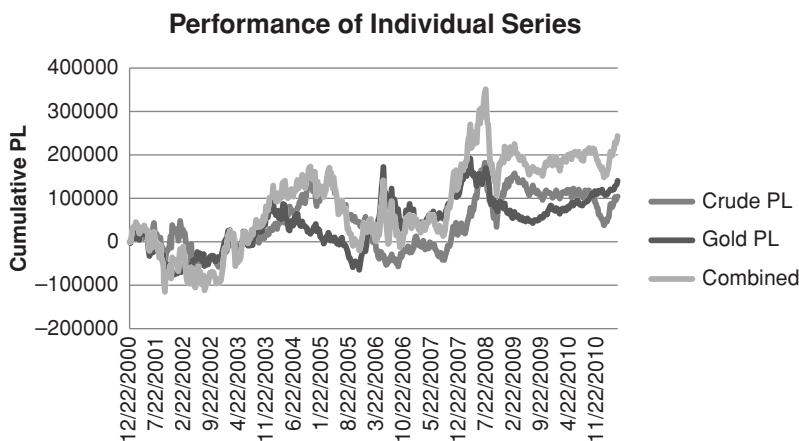
(b)

FIGURE 13.17   (a) Performance of the crude/gold ratio, and
(b) Performance of the individual crude and gold contracts with the
combined performance.

trades much more often, while the entry price for both crude and gold stay the same for
long periods. This offers some hints about changing the sensitivity of the ratio method
to reduce the number of new positions. It also shows that the two methods are very dif-
ferent and, as we expected, create a new market that can offer diversification to a trend-
following portfolio.

There are many markets dependent on the price of crude oil; energy, plastics, rubber,
and transportation companies lead the list, although oil is part of synthetic material and
many other products. If we apply the same 40-day moving average to the ratio of Exxon-
Mobile (XOM) to crude oil, this time volatility adjusting using the average true range
method discussed in the section "Pairs Trading," with position sizes based on a $10,000
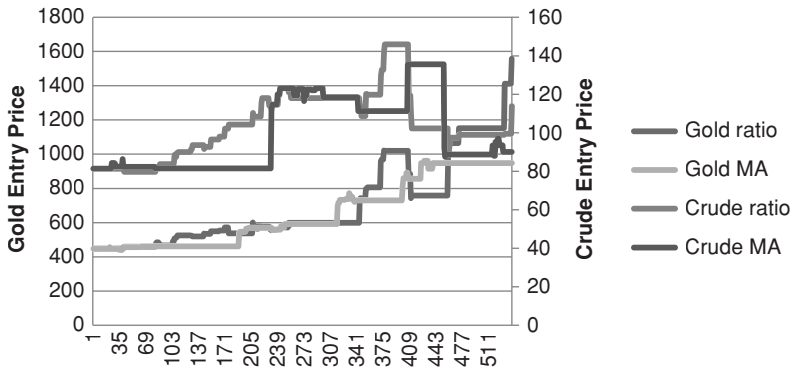
FIGURE 13.18   Comparison of crude and gold entry prices using the ratio method and individual moving averages.

investment in each leg, we get the results shown in Figure 13.19. In this case, each trade had the same risk assigned to each leg. Only the closed-out cumulative profits are shown. While crude shows two nasty losses in the middle of the chart, Exxon shows a large loss at the end, offset by profits in crude. You may choose to trade only those positions setting longs in XOM, although that stock is readily available for short sales.

In the last example, one of the four largest trucking companies, Werner Enterprises (WERN), is matched with crude oil, expecting a drop in the share price of Werner when crude moves higher (the ratio declines). In this case, the 40-day trend seemed to work for both markets using the ratio as the basis for trading signals, and the results are smoother due to the diversification provided by the two markets. The position sizes of both
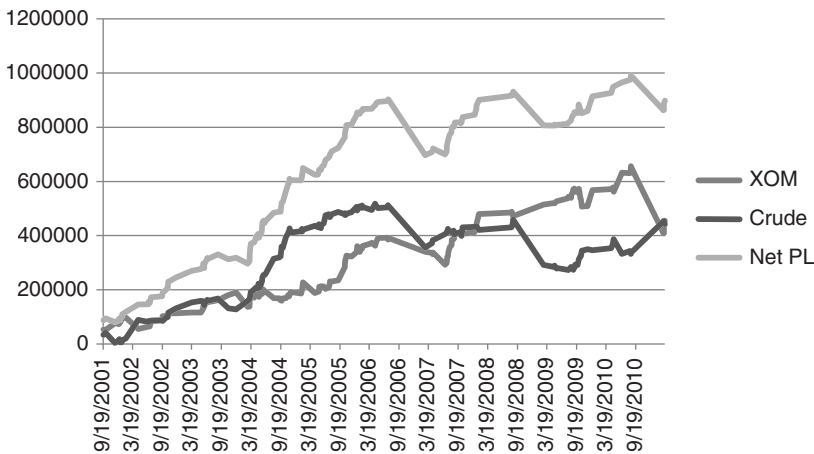
FIGURE 13.19   Trading the Exxon-Mobile/crude oil ratio using volatility-adjusted position sizes.
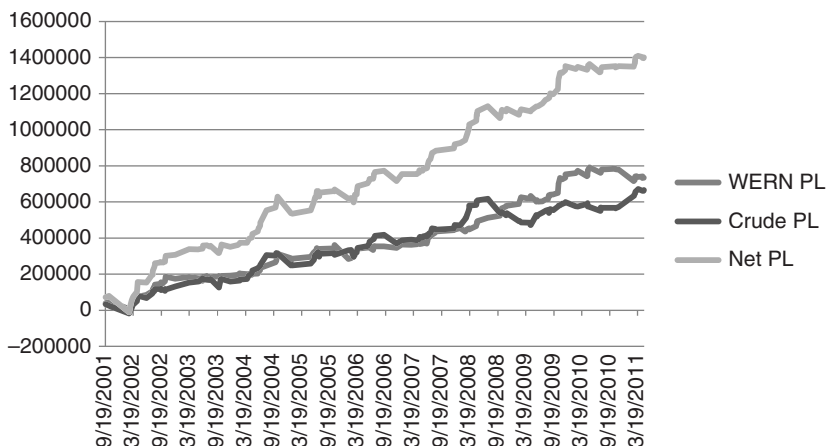
FIGURE 13.20   Results of trading the Werner Enterprises (WERN)/crude oil
ratio, daily data.

markets were adjusted for volatility to equalize risk. Figure 13.20 shows recent periods
where profits and losses in one market were offset by the opposite performance in the
other market, netting a smoother combined result.

The components of the ratio spreads were chosen because it was believed that
they had a fundamental relationship. Those choices all seemed to work. Other combi-
nations would be commodities/gold, crude oil, and the euro currency; however, there
may be many more. There was no comparison of whether these results were better
than trading the individual legs using the 40-day or 40-week moving average because
the purpose was to create a new market and generate trading signals at different times.
In the world of hedge funds, liquidity can be a problem, and this method helps to solve
that problem.

Trending or Mean Reverting?

The previous sections have looked at spreads in many ways, using pairs, ratios, old and
new crop, and trends. There have been example of intermarket and cross-market spreads.
How do you know whether the spread strategy should be trending or mean reverting?

At different places throughout this book, we've discussed the existence and effects
of market noise. Noise is most noticeable in the short term, and trends overwhelm the
noise in the long term; therefore, when we choose to use a mean-reverting system, we are
trying to take advantage of market noise and short holding periods. On the other hand, if
there is a sustained divergence in two stocks, then a longer-term trend will capture that
effect best.

By its nature, pairs trading seeks profits from short-term price distortions; therefore,
it succeeds in the short term. When investors shift from high-risk stocks, such as low
caps, to more secure dividend-paying stocks, it is a gradual move of sentiment that can
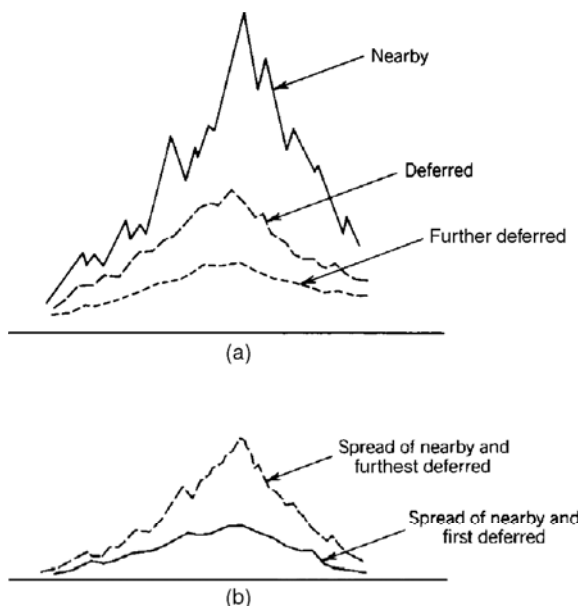
FIGURE 13.21 Interdelivery spread volatility.
(a) Actual prices. (b) Relationship of interdelivery
spreads.

be followed using a trend. Similarly, when interest rates fall, investors tend to move to shorter-term maturities so that they are able to move their money when interest rates start rising.

Trend Following Using Bull and Bear Spreads

A trending spread can be a lower-risk substitute for an outright long or short trend position. In grains and foods (commodities in general), a bull market is the result of prices of the nearby delivery months rising faster than deferred months. A *bull spread* can be placed by entering a long position in the nearby and a short position in a deferred month. Both risk and reward are reduced; however, the greater the time between months, the more volatile the spread (Figure 13.21). As in an outright position, once the upwards price peaks, the spread must be reversed because the nearby delivery will decline faster than the deferred. A *bear spread* is entered by selling the nearby and buying the deferred contract.

The analysis that identifies the time to enter a bull or bear spread can be a standard trend-following approach based on the nearby contract only. Many traders believe that the spread itself must also confirm the trend before a position is taken; therefore, a medium-speed moving average may be applied to the spread series. This would provide a signal based on the relative change in spread direction. A 1- or 3-day change might be enough for the fast trader. A bullish trend signal in the nearby delivery and a bearish signal in the spread create a conflicting message, and no trade should be entered.
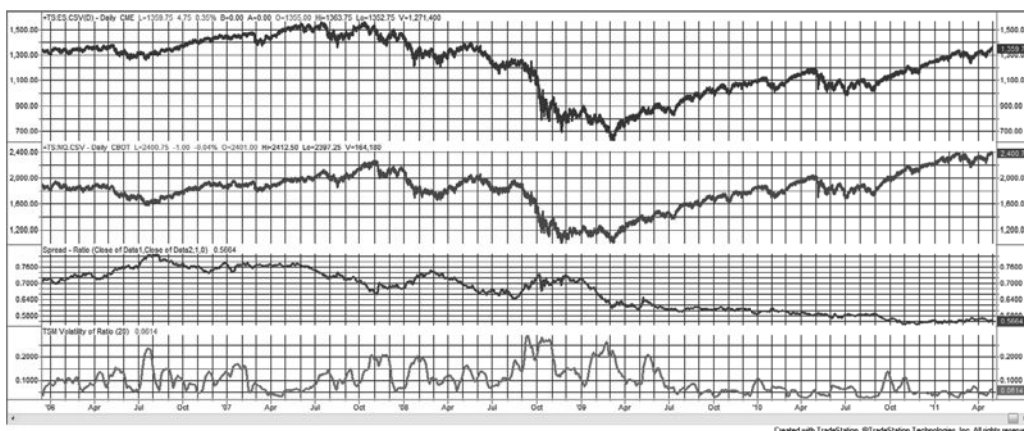
FIGURE 13.22   Annualized volatility of S&P/NASDAQ price ratio, using futures, shows three different levels of volatility over five years.

Changing Volatility in Spreads

The use of relative-value spreads means that you are looking for situations where one market and/or stock is relatively high and the other is relatively low. That generates a trading signal to sell the highest and buy the lowest. Because it's a matter of relative value, we know that the timing is good, but we do not know the volatility of the two legs. If that volatility is very low, then there may not be enough profit potential in the trade.

Volatility of the two legs, taken individually or together, can create either high risk or no opportunity for a mean-reverting strategy. Occasionally, jumps in volatility can create large losses, and most spreads suffer this problem. Figure 13.22 shows the electronic S&P and NASDAQ futures price from 2006 into 2011. The ratio is shown in panel 3, and the annualized volatility of the ratio of these two markets at the bottom. If you were using a mean reversion approach to trading based on the actual ratio values, then buying at a ratio of 0.05 and selling at 0.15 during the first half of the chart would have been good. However, during the crisis that started in September 2008, selling at 0.15 would have meant suffering a catastrophic risk. The period that followed starting in July 2009 is just as much of a concern. Volatility falls and the ratio no longer touches 0.15.

This all argues for relative-value indicators to time entries and exits. Still, the jump from the earlier levels of volatility to the extremes of 2008 will produce a loss using any method. And while relative value will produce trading signals during the low-volatility periods from mid-2009, it will not tell you whether volatility is too low to generate a trading profit because returns are less than the trading costs. To overcome that problem, a low-volatility filter is needed, which will reduce the number of trades even at the same time that it selects trades with more profit potential.

Historical Comparisons

Analysis of longer historic periods will result in a better understanding of the consistency and risk involved in trading the spread. A long-term study will reveal seasonal and cyclic

moves as well as periods of lower and higher volatility. Some of the factors that affect changing volatility patterns are:

1. *Underlying price-volatility relationship.* Higher prices in most markets are associated with higher volatility and wider spread differences; lower prices force most spreads to narrow. The spread between heating oil and gasoline must be greater when those products are trading at $4.00 per gallon than when they are at $1.00. In addition, volatility increases with sudden price changes, but when the public adjusts to those levels, even if they are much higher, volatility declines. It is the concept that volatility increases during uncertainty then declines at equilibrium.

2. *Normal seasonal patterns.* Seasonality causes spreads to narrow and widen in predictable ways. During harvest, the *basis spread* (the difference between the spot and cash markets) widens, reflecting the available supply. After harvest, the basis will continue to narrow if demand fails to materialize.

3. *Investor risk preference.* Investors shift to safer holdings during times of uncertainty and then return to markets with higher risk as the economy returns to normal. One familiar pattern is between the stock market and interest rates, where the phrase "flight to quality" indicates the move to government interest rates when a destabilizing event occurs, either a large drop in the stock prices or a geopolitical crisis. There is also a shift between different segments of the stock market. During good economic periods investors are willing to risk the uncertainty of small caps; otherwise, they find safety in the larger companies. Figure 13.22 compares the last five years of the S&P 500 and NASDAQ, where NASDAQ is considered the riskier investment. Prior to 2008, ratios were much higher, indicating a larger portion of investors were putting money into NASDAQ. From 2009 that ratio dropped more than 25%, showing significantly less interest in riskier assets.

4. *Specific events.*
   a. A freeze in coffee or orange juice that reduces supply and damages trees has a prolonged effect on prices and will make prices in the following crop year even more sensitive to a potential freeze.

   b. Events that cause seasonal price changes, such as exceptional demand for exported grain, will overwhelm the seasonal patterns for the remainder of the crop year and often into the next season.

   c. A combination of circumstances, such as low inventory, higher export demand in the previous year, and dry, hot weather midway in the growing season, will result in a nervous market, causing prices to rally for the smallest reasons in anticipation of potential problems becoming realities. By observing past similar situations, the extent of the rally can be seen as large enough to use a medium to fast trend-following method, with profit-taking either at the time the trend turns or 1 month before harvest begins, whichever comes first.

Timing the Relative-Value Trade

In the section on pairs trading, the fast stochastic was used to identify the timing for entries and exits. That is still a viable alternative, along with many other oscillators such as the RSI and MACD. One newer calculation, Ehlers' *Fisher Transform* is particularly interesting (the formula can be found in Chapter 11). Although it starts out to be similar to the stochastic indicators, it is structured to avoid getting stuck at the highs and lows for extended time periods. The full range of the Fisher Transform is +1 to −1 (or +100 to −100), but overbought and oversold levels at about +80 and −80 are recommended. When the ratio peaks above 80 (see Figure 13.23) then platinum would be sold and gold bought. The trade would be closed out when the ratio is near zero, indicating a return to normal. Ehlers' Fisher Transform does a good job of moving quickly between overbought and oversold.

Rating Services

The Internet is a source of many financial services that rate the likely success or failure of stocks during the next few weeks or months. Some of these companies have performance records to show that their rankings have been reasonably accurate. The most notable of these is *Starmine*. These services can also charge a substantial fee.

Alternatively, there are websites that allow you to rank stocks by various criteria, such as the price/earnings ratio (P/E). These include *TheStreet.com*, Cramer's *Mad Money*, *Reuters*, and *The Motley Fool*. Other websites are closer to what is found on Starmine by giving a summary of analyst recommendations. These take the form of *strong buy*, *moderate buy*, down to *strong sell*, each with numeric values associated with the level of rating. Whether you create your own ranking or use the analysts', it is necessary to record the daily or weekly (preferable because the results might be noisy) rankings in a spreadsheet and monitor their success.
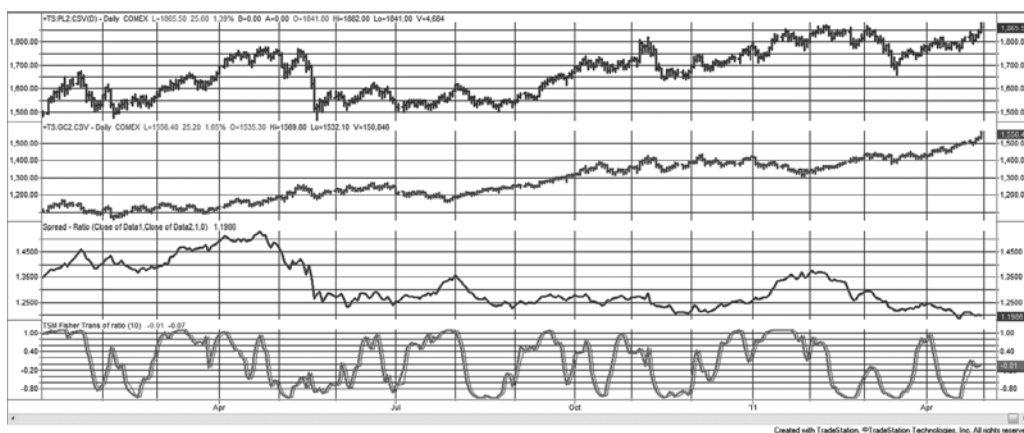
FIGURE 13.23 The platinum/gold ratio using futures (third panel) with the Fisher Transform at the bottom.

Having established whether the service or your own method is good, there are ways to trade the results as spreads:

1. If the total number of stocks being ranked is small, then trade between 20% and 35% of the top and bottom stocks in the ranking. If the number is large, then you need to trade as many as offer diversification, at least 5 from each end, but more is better. The stocks chosen at the top and bottom will be called *baskets*.

2. If the ranking is successful, then you will be treating it as a trending method, buying the best and selling short the worse. If the ranking is bad, then it will be mean reverting, and you will sell the best and buy the worst.

3. If you are trading n number of stocks in each basket, and stock n (the worst of the best or the best of the worse) moves out of its basket by at least two positions, then close out that position and replace it with the one that has moved into that basket. This way you avoid frequent switching.

4. Verify that the stocks in the bottom basket are moving and have volatility. In some cases stocks are ranked very low because they are rarely traded. Those stocks will not offer any risk protection and will distort the two baskets of stocks. The top basket should not have this problem.

5. Volatility-adjust each basket, preferably using the dollar value average true range and equal investments for each stock, in the manner shown in the "Pairs Trading" section, earlier in this chapter. Similarly, the baskets may diverge in terms of volatility and would need to be adjusted, but that should only be done if they vary by more than 20%. Otherwise, each new position will be adjusted for current volatility, and the portfolio will automatically self-adjust.

Self-ranking can be done many ways, from using fundamental data about the individual companies, or simply by tracking the returns, or the information ratio, the returns divided by the risk, or even by the performance of a long-term trend system. For fundamental analysis, the ratings available on the Internet seem best, but for technical rating, you will have more control over the process if you do it yourself.

Generalized Market Neutral Strategies

The classic market neutral program is a mean-reversion (arbitrage) strategy applied to an individual industrial group, or a correlated set of stocks. Generally, the correlations should fall in the range of 0.50 to 0.70. Anything lower exhibits more risk, and anything higher has difficulty producing enough profits to offset trading costs. As with pairs trading, cointegration may be a better way to choose participants than correlation. One advantage of larger market neutral programs, those with many stocks traded simultaneously, is that some participants with lower correlations may add to the larger per trade results but have their risk offset by diversification.

During the past 10 years the propagation of index funds and program trading (arbitraging individual stocks against the index) has caused many stocks to have increased correlations. When the S&P futures drift away from the cash S&P due to orders from large hedge funds or insurance programs, and that drift is large enough for a profit, electronic orders will be generated instantly to buy all the stocks in the S&P 500 and sell the futures. The act of buying all the stocks increases the correlations and reduces the opportunities for successful market neutral trading.

Essentially, the market neutral strategy is very similar to the steps shown in the previous section on rating services. The overall level of price noise in the stock market, however, is very high, which makes a mean-reversion strategy more likely to succeed than a trending approach; in which case, a short holding period is better than a long one to avoid fighting the underlying trend of the market as a whole.

Other Spread Issues

Experience shows that some trading techniques are better than others, and some are dangerous.

Using Conditional and Spread Orders

The opportunities for entering a spread at a good differential can be short-lived and require a spread order to ensure that the minimum objectives are met. Instead of entering two orders to buy stock A at a specific price and sell short stock B (or an index) at another price, it is far better to place a spread order to buy A and sell B at a difference of x. It is a common order for futures markets, but more difficult for stocks because one leg is a short sale.

Alternatively, the short sale should be done first; then the buy follows as soon as the short sale is confirmed. Some trading platforms show when there is stock available for short sales, so placing the two orders simultaneously by pushing the enter key would be the simplest method. In futures, spread orders have the advantage of being cheaper, about three-quarters of the cost of two outright orders.

Legging In and Legging Out of a Spread

When both *legs* of a spread are not entered or closed out at the same moment, trading risk is increased. An unprotected leg of a spread is simply an outright long or short position and must be managed. Consider a bull spread in a fast-moving market, where the nearby futures price is rising faster than the deferred. Although both legs can be entered simultaneously using a spread order, it may be tempting to enter the long position (the price moving in the trend direction) first, followed by the short leg within a few minutes or a few days. The intention is to take advantage of the price direction. Ideally, the short is entered after there are profits on the long side, and those profits are then protected. Legging in is a rewarding philosophy when it works. When it doesn't work, there is a large loss on the outright position, which will be difficult to offset with the lower profits from the spread. If the purpose of trading a spread is to reduce risk, legging in or legging out is not going to accomplish that goal.

When the trend turns from up to down, it is necessary to switch the spread from bull to bear. If large profits have been made in the bull move and both the trend and spread signal a downward turn, it could be tempting to lift the long leg, and hold the outright short using accrued profits to offset the increased risk. A more conservative trader might enter a bear spread in the two delivery months that represent lower volatility thereby reducing risk and profits even further—thereby conserving prior gains. The volatility of some futures markets such as energy are lower in deferred months, while others, such as Eurodollars, are higher.

Creating a Spread to Protect an Outright Position

Protecting an existing trend position by spreading with the next deferred delivery, or by using options for either stocks or futures, is a questionable strategy. If the market has been going up and indicates a temporary downturn, some traders will want to *hedge* by selling the next deferred contract, an ETF, or an option in the case of stocks.

For stocks this method is justified because of tax advantages that result from a long-term capital gain. In other cases, there is an added cost for hedging, commissions or option premium, plus the added complexity of the new position. Not only is it necessary to track the original trade, but you will need to know when to remove the hedge. Given the competitive commissions costs these days, closing out a trade may be a simpler and more effective way to deal with uncertainty.

Reverse Response to Trending Markets

Some intradelivery spreads respond in just the opposite way to trends. Normally, rising futures prices are reflected most in the nearby contracts and less in the deferred contracts, where there is less certainty. For precious metals, or pure carry markets, an upward trend results in the deferred contracts rising faster than the nearby. For example, rising gold prices result in a larger total contract value; consequently, there is higher interest due on the contract value, a component of the carrying charges. Each deferred contract will reflect increased carrying charges.

Exceptions to the Rules

During a trending period, most markets exhibit a clear relationship between delivery months. However, this is not the case for nonstorable commodities, such as cattle, hogs, broilers, and eggs. These markets can show little relationship in the response of deferred contracts to the same bullish or bearish news.

During periods of exceptional demand, or if feed prices are very high, livestock may be sent to market early. This causes a drop in current prices and will force the price of hogs lower as well. The price of cattle and hogs in the next deferred month may rise to reflect the shortage of supply; however, even in the most extreme cases, this pattern cannot be carried very far into the future.

Extreme Spread Ratios

The two sides of a spread rarely have the same volatility. When badly mismatched, a spread will act the same as if you had taken an outright position in the more volatile leg
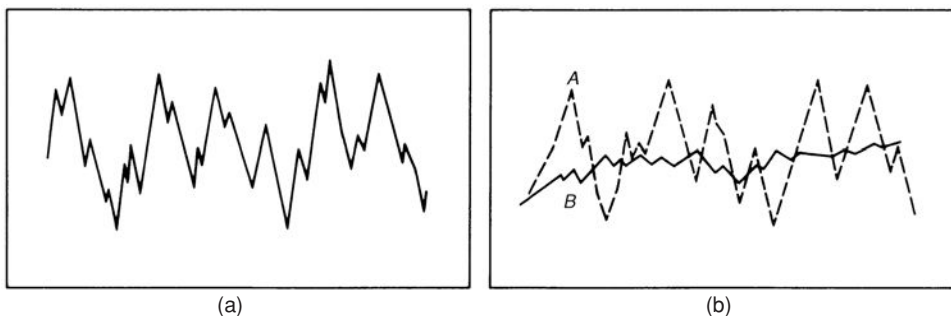
FIGURE 13.24   Poor spread selection. (a) Spread price. (b) Spread components.

of the spread. The risk of the spread trade may even be higher than an outright long or short position. For example, Figure 13.24a shows the price movements of a spread, and Figure 13.24b shows the components. Leg A is the most volatile. In the case of a futures market spread using delivery months of the same product, the deferred months are progressively less volatile, with the exception of interest rates. This is often the case because news that affects the current price has less of an impact on future prices. A volatility mismatch, as shown in Figure 13.24b, may be a poor choice for a spread, even if the position sizes are adjusted for the difference. Given one stock trading at $100 and the other at $2, we may have a liquidity problem in the cheaper stock. If you trade 1000 shares of the expensive stock you would need 50,000 shares of the cheap stock. Extreme ratios distort the importance of a move in the cheaper stock. Small moves that are not based on anything significant will be erratic, causing higher overall risk. In addition, low-priced stocks tend to have higher relative volatility even though it has lower absolute volatility.

The large position size presents event risk in the cheaper stock. Low-priced stocks can make much larger moves than higher-priced stocks. Even at $5, good news in a stock could jump the price to $5.50, a 10% move. A 10% move in a $100 stock will be a much less likely. When leveraged up in the cheap stock, a move of 10% in the wrong direction could be fatal.

In futures, there are other factors affecting the legs of a spread.

- Specific events, such as an economic report or Fed announcement, can affect legs differently. Fed interest rate policy usually targets different ends of the yield curve, affecting the volatility in different ways along the curve.
- Highly volatile spreads in nearby months may not adjust to normal before the expiration of the contract. At expiration, extreme demand in the cash market may cause further unusual spread differences cause by a *short squeeze*, the inability to deliver the cash product at futures expiration.
- A highly volatile spread may have greater risk than a single outright position in one leg. This can be the result of the two legs moving in different directions or the added leverage gained due to lower margin. This is not the purpose of a spread trade.

Behavioral Techniques

Some approaches to trading are directly dependent on human behavior and cannot be represented by pure mathematical techniques. Short-term systems are more likely to target investor behavior than economic factors because, over a few hours of one day, the influence of macroeconomic policy and long-term trends is very small. The concepts presented in this chapter deal specifically with the interpretation of human reactions, although the interpretations are all systematic. The section "Measuring the News" covers an area that has always been important to a floor trader, yet has been greatly overlooked by technicians; it offers great opportunity for research. "Event Trading," the study of the market's reaction to price shocks, has become increasingly important for both proactive and reactive traders. "Opinion and Contrary Opinion" takes the form of a poll or consensus of opinions of traders and market publications. It may help answer the question, "What is everyone else doing?"—or at least—"What are they thinking of doing?"

The principal works of Elliott and Gann are included in this chapter with a complete discussion of the Fibonacci series and its ratios. Fibonacci forms a singular part of their techniques, and has been applied to most other forms of charting. The way in which traders respond to market moves and the remarkable similarity that can be found in Nature give serious underlying substance to these methods. Because not all of the assumptions upon which these systems are based can be quantified, they can only be substantiated by the performance of the systems themselves. All of these behavioral methods are fascinating and open areas of creativity essential to broadening system development. They are grouped together with discussions of natural phenomena and the rapidly growing area of financial astrology, all of which should stimulate your grey matter.

MEASURING THE NEWS

If you can keep your head when all about you are losing theirs, maybe you haven't heard the news.

—Rudyard Kipling, adapted by H.L. Menken[1]

The news is likely to be the greatest single element affecting pricing of free markets. It broadcasts both official statistics and interpretation, and can express information directly and by innuendo. It is indispensable to traders. If the news services are not objective, they could materially alter any opinion by including unverified claims or omitting relevant information. Occasionally, one news service releases an economic report too soon, throwing the market into a turmoil. On the whole, financial reporting is good, but cannot be without bias. The very act of interviewing the CEO of some company may give viewers the impression that the company has credibility, which is not always the case.

The impact of news is so great that a speculator holding a market position according to a purely technical system would do best not reading, listening to, watching, or in any sense being exposed to news, because whether pro or con, it is impossible to ignore. In a study commissioned by the *Wall Street Journal*, it was shown that 99% of the financial analysts polled read the paper regularly, and 92% considered it the *most valuable* publication they read.[2]

As an element of a trading program or as an indicator of its own, the news has exceptional value, but its interpretation can be very complex. We believe that, if we could measure the impact of unexpected news, the importance of the *Wall Street Journal* or other wire service articles, earnings releases, economic reports, the USDA crop reports, and the CFTC Commitment of Traders report,[3] we would be able to anticipation the direction of prices, or at least react profitably. But first we must be aware of the complications of analyzing the news. There is the problem of objectively selecting which items are relevant and which are most important. And then there is the difficulty of quantifying the news—how do you rank each item that is measured, and on what scale do you determine cumulative importance? Some news items clearly have a greater impact on the market— weather disasters, terrorist activity, economic reports, a change of interest rates, major trade agreements, changes in government policy, key crop reports—but these must be ranked as bullish (+) or bearish (−) in a way that produces a numerical system of analysis. On a single day, an address by the president about foreign trade may be ranked −3, a meeting of the Federal Reserve +7, unemployment −4, a bombing in the Middle East −3,

---

[1]Adam Smith, *The Money Game* (New York: Random House, 1967). Based on the poem "If" by Rudyard Kipling.

[2]Frederick C. Klein and John A. Prestbo, *News and the Market* (Chicago: Henry Regnery, 1974), 3.

[3]A thorough analysis of the CFTC *Commitment of Traders* report can be found later in this chapter. It should be noted that other government reports released on the same day may complicate the interpretation.

a continued lack of rain in the West +2, larger-than-expected retail inventories −4, and a key article in the *Wall Street Journal* on the improved Libyan crude oil production +5, giving a net score of −5 to the overall economic picture. This would be interpreted as moderately bearish.

Klein and Prestbo attempted such a study by assigning values of 3, 2, and 1 to articles in the *Wall Street Journal* of decreasing importance. Their interest was the stock market, and their work was straightforward and some of the conclusions general in nature. They showed a direct correlation between the relevant positive and negative news articles and the direction of the stock market. As it was scored over 6-week intervals before and after major turning points in the Dow Jones Industrial Averages, the news would stay about 70% favoring the current market direction. Having eliminated the possibility of the market influencing the news, they could conclude that, in retrospect, the market reflected the nature of the news. A victory for common sense.

With the continued globalization of world futures markets, there is a noticeable reaction in all financial markets to significant economic reports released by any major economy. The correlations are especially high during a crisis when markets are under stress. Rising consumer confidence in Germany is interpreted as bullish in the United States because it is positive for American exporters. Surprisingly lower earnings or disappointing growth of a major microchip manufacturer reflects badly on the entire tech sector everywhere in the world. News of the need to restructure Greek or Italian debt will cause a movement away from the euro, while the resolution will strengthen the euro.

There are exceptions when news is of limited scope and relevant to a specific commodity, such as a lack of rain in the Corn Belt during prime growing season, or a winter freeze in Florida's orange orchards or Brazil's coffee region. The news is watched carefully during the last phase of a bull or bear move. Commodity traders will wait for a crop report to show the final numbers on yield, in the same way the financial market traders will wait for initial claims to fall or the Fed or European Central Bank to raise rates, signaling a final change in the economic trend. Chicago weather traders are notorious for watching news that affects U.S. crops. It is the anticipation of weather, rather than the actual weather, that drives most price moves, as it is the anticipation of interest rate changes, or employment numbers, that moves the markets. The news items that affect prices the most are:

- Release of economic statistics (PPI and CPI for inflation, retail sales, balance of trade, employment and initial claims, consumer confidence, housing and refinancing).
- Action by the Federal Reserve or other central banks to change rates or change their *bias* (the way they express future policy).
- Changes in the money supply that indicate easing or tightening.
- Government reports on commodities production and inventories.
- Unexpected news or price shocks, such as an assassination or terrorist attack
- Trade negotiations, agreements, legislation, and occasionally rulings by the U.S. Supreme Court that affect business.

- Weather and natural disasters, such as the Japanese earthquake, or a hurricane that affects the sugar crop or causes large payouts by insurance companies.
- In-depth studies by the *Wall Street Journal.*
- Front-page news articles and dominant television coverage of high prices, strikes, etc., and their potential effects on business.
- Market letters, research reports, and comments from accepted authorities, major brokerage houses, and influential organizations.

Ranking and Measuring

The problem of ranking and assigning numeric value to news items is that it requires knowledge of how others see the news. Klein and Prestbo studied this problem for the stock market and concluded that about 90% of the *Wall Street Journal* readers perceived news in the same way, when classified as bullish, bearish, or neutral. A reason for the consistent interpretation of the news articles is the published analysis. Within minutes of the release of an economic report or Central Bank action, the financial news services broadcast expert opinions. When this study was done, interpretations were supplied by the brokerage houses and transmitted to the media to be relayed to traders. The *professional analysis*, then and now, is assumed to be correct and is an overwhelming influence shaping public opinion. The report of an upgrade of Ford from *hold* to *buy* by Bank of America will bias investors towards being long Ford.

News can also be measured empirically, by studying the immediate impact of an expected or surprise news item. For statistical reports, such as Initial Claims, Consumer Confidence, or the Producer Price Index, care must be taken to use the correct figures. Market reaction is a combination of expectations compared with actual figures and corrections to the previously released data. In some cases, it is difficult to know whether a jump in prices was due to an upwards revision of the previous data or current values that are higher than expectations. You can find this week's list of economic reports on the Internet, along with the previous report data, expectations of the economists for the new data, and expectations of the market. Logical Information Machines (LIM), a data service, provides all of the historic reports including revisions, and provides tools for analyzing the impacts. Combined with a record of trader expectations, there is enough information to evaluate trader reactions. Following the release of an economic report, we can then expect the price change to be expressed as:

Price change $= a \times ($ *Current value – Expectations* $) + b \times ($ *Revision – Previous value* $)$

where a and b are weighting factors, $a > b$ implying that the current value is more important than the previous one. The values of a and b can be found using a regression analysis. It is necessary to make the assumption for this type of measurement that the effects of a news release are most important in the short term, and that their influence on the market is diluted each day (and often each hour) following the release. A starting

point for representing the way in which the news decreases in importance is to use the relationship in physics:

$$I = \frac{1}{T^2}$$

where I = the net impact
             T = the elapsed time since the release of the news

In science this relationship, which represents the physical impact declining with the square of the distance or time, applies directly to sound and light.

The impacts of economic reports are significant, frequent, and worth studying. For futures markets, the CFTC releases its *Commitments of Traders* (COT) report each week. It tells the distribution of holdings among large and small speculators and hedgers as a percent of total open interest. This report is watched with the idea that the small speculators are usually holding the wrong positions and the positions of large hedgers are normally right. The COT report is covered in a later section of this chapter.

Trading on the News

Even without a sophisticated method of measurement, there are many professional speculators who trade on the news. When a bullish news item appears and the market fails to respond upwards, the experienced trader looks for a place to sell. It shows that expectations exceeded reality and prices had already anticipated the bullish interpretation of the news. Sometimes there are a large number of sell orders placed above the market in anticipation of one last rally after the report is released. An increasing number of traders watch the early morning price movement on 24-hour electronic exchanges, before the primary market opens. The *primary market* is the local stock market or the beginning of the normal business day. One strategy is to sell a higher overnight move when there is little news to support the rise. If prices have been rising steadily in the premarket session in anticipation of a bullish economic report, selling early may be the only opportunity. Prices frequently gap lower after a bullish report that did not meet market expectations. Most U.S. economic reports are released at 8:30 A.M., New York time, one hour before the U.S. stock markets open. The response to the report can be seen on Globex, where most major index markets are traded. If you are trading individual stocks or ETFs then you may be able to trade through an ECN (Electronic Communication Network), the first of which was Instinet in 1969. Since then, the most successful, Island and Archipelago (Arca) have been absorbed by NASDAQ and the New York Stock Exchange, respectively. These off-exchange providers allow you to trade at any time and match your order with others within their own network. Good executions depend on finding other orders at your price. You may also find access to off-market trading through the major electronic order entry platforms, Think or Swim (TD AmeriTrade), Schwab, E*Trade, and Interactive Brokers.

Agricultural *weather markets* function purely on news. When there is a potential for a drought, traders with long positions wait for the 5-day forecast hoping for no rain; they anticipate a loss of a specific number of bushels per acre for every dry day once rainfall is below a specific level. Weather markets are nervous, with prices often gapping higher and lower at the open, and are characterized by evening-up on weekends; they rely heavily on anticipation. It is said that a farmer loses his crop three times each year, once for drought, once for disease, and once for frost. In 1976, the news carried numerous articles on the desperate wheat crop in the western states, showing films of virtually barren fields, and yet the United States harvested one of its largest wheat crops on record. Weather markets have a history of volatile price movement but rarely materialize as the disaster that is anticipated, thanks to new drought-resistant strains of hybrid corn.

The *discounting of news* is as important as the news itself. An old saying in the market, *buy the rumor, sell the fact*, implies that anticipation drives the price past the point where it would realistically account for the news. When the actual figures are released, there is invariably an adjustment back to their proper level. The pattern of anticipation for each economic report or news event should be watched closely. A later section in this chapter, "Trading the Reaction of Treasuries to Economic Reports," shows that price reversals are more common than continued movement following a report.

Market Selectivity

The market seems to focus on one news item at a time and one remedy at a time. Although the same factors are always there to affect prices, they must reach a point of newsworthiness before they become the primary driving force. For heating oil, the combination of unexpected, sustained cold weather compared to inventories will activate a weather market. Although professionals may monitor thermals, published by the U.S. Weather Service, the market will react quickly to a weather report that anticipates a cold front over the next five days.

Since 2003, when the economic recovery became the leading driver of stock prices, U.S. traders have focused their attention on weekly Initial Claims and the monthly employment report. The news broadcasts repeatedly stated that 400,000 claims was the balancing point between job growth and job loss. A sustained economic recovery needed job growth which would translate into consumer spending. The weekly Initial Claims data, released each Thursday at 8:30 A.M. in New York, is still followed by a sharp reaction in Treasuries and Index futures in U.S. and foreign markets. After years of terrorist attacks, wars in Iraq and Palestine, political upheaval in the mid-East and northern Africa, and the subprime crisis, the market still responds to Initial Claims and ignores most other news. It is a classic case of crowd behavior.

Media Indicators

In a delightful article,[4] Grant Noble argues that the news recognizes events when they are cresting, and most often provides a countertrend trading opportunity. First, the

---

[4]Grant Noble, "The Best Trading Indicator—The Media," *Technical Analysis of Stocks & Commodities* (October 1989).

American media should be viewed as providing trading signals in three major time frames:

1. *Long term*, as given by the large circulation news magazines such as *Time*, *News-week*, and the *Economist*. With the timeliness of a brontosaurus, they profile moves that last many years.
2. *Medium term*, represented by *Barron's*, *Forbes*, and *BusinessWeek*, covering a period of about 3 months.
3. *Short term*, held captive by the *Wall Street Journal* and the *New York Times*, which provide intermediate predictions as well as medium-term outlook.

By reading these periodicals you find that the *Wall Street Journal* has run headlines on a "killing drought," "dust bowl," and the *New York Times* on "Drought…imperils crucial wheat crop" just as the wheat price makes its highs for the current move. In another case, *Barron's* cover article asked, "Is the bull leaving you behind?" in August 1987 just ahead of the precipitous drop in October of that year. It is not surprising that the media would highlight events only after they have become a popular concern. We might say that the proportion of news coverage corresponds to a high public consensus, a topic addressed in the section "Contrary Opinion." It may be perfectly valid to construct a consensus indicator based on the number of square inches of news coverage given to an event in a combination of publications, each weighted by its circulation. This may be easier now by scanning the most popular electronic news websites. Even with the increased access of electronic news, broad coverage and recognition of consensus remains a lagging indicator.

EVENT TRADING

There are two situations that trigger *event trading*. The most common are scheduled economic reports, earnings announcements, and recommendations by financial institutions and independent analysts. The price moves tend to be modest, and traders usually take a position opposite to the market reaction, looking for a correction. This was an area discussed as part of arbitrage in the previous chapter. In this section we will look at larger reactions to reports and unexpected events, and address specific situations.

The largest price moves with the greatest volatility, called *price shocks*, are the result of reactions to perceived important, unexpected news. These market events pose the greatest risks to all traders because they are unpredictable and of such great magnitude that they are out of proportion with normal trading risk. It may be possible to trade for a number of years without experiencing a large, adverse price shock; therefore, many traders do not plan properly for these situations. Yet surviving a price shock is often the difference between the end of a trading career and a long, successful one. Further discussion of price shocks can be found in Chapters 21 and 22. This section presents a strategy for trading immediately after a price shock.

Not all price shocks are of such magnitude that they present an unmanageable problem. It may be that most price moves, big and small, are either the reaction to news or the anticipation of news. Larger price moves are the result of an act that affects the global economy, such as a terrorist attack, declaration of war, assassination, or natural disaster associated with weather or earthquakes. Smaller shocks come from the periodic release of economic reports by a government agency or monetary authority, financial commentary on television, radio or newspaper, earnings releases, unexpected bankruptcy, rulings by the Supreme Court, statements by the Department of Agriculture with respect to a new drug, and countless other sources. The U.S. government releases economic data on a preannounced schedule, many of them at 8:30 A.M. New York time. These create regular disruptions.

The frequency and size of these price moves, triggered by unexpected news, make these events a natural candidate for a trading system.[5] The profit opportunity, however, does not lie in taking a position ahead of the report and anticipating the market reaction, but in studying the systematic patterns that come after the initial price reaction to the news. Because the news is unexpected, you cannot predict the results or the extent of the reaction; therefore, taking a market position in advance of a government report would have a 50% chance of success and often very high risk. Studies might show that there is a bias in the direction of the price shock due to the way the monetary authority plans economic growth and controls inflation; however, the risks would remain high. This section only looks at positions entered on the close of the *event day*, the day on which a sharp market move occurred.

Market Reactions to Reports

To determine whether there is ample opportunity to profit from the price move that follows an *event*, it is necessary to study the direction of the price changes that occur over the following few days. These lagged reactions are the results of market inefficiencies. When a surprise occurs, no one can know what the correct price should be; it takes time for the market to reach equilibrium. The premise proposed by the Efficient Market Hypothesis, that the market will immediately find the correct price, is ridiculous. It may the correct price at the moment, given the time to react and the size of the orders flowing into the market. With large price shocks there is often an over- or underreaction that is corrected during the next few days. Sometimes prices jump one direction and immediately begin to go the other way until they have completely discounted the price shock by the end of the trading day. The initial reaction to a price shock may not be the profitable direction for a trader.

Figure 14.1 illustrates the types of price movements that are expected follow an upwards price shock. When there is an underreaction, or a structural change, prices move higher over the next few days; when there is an overreaction, or a short-lived event,

---

[5]Ben Warwick, *Event Trading* (Chicago: Irwin, 1996). Also see Andrew Busch, *World Event Trading* (Hoboken, NJ: John Wiley & Sons, 2007) and Peter McKenna, *The Event-Trading Phenomena* (Wilkes-Barre, PA: TradeWins, 2003).

prices reverse. Patterns are not as orderly as shown in Figure 14.1 because volatility is high and many traders react impulsively (or by financial necessity) to the move, covering their losses. A move that is a candidate for event trading:

- Must have enough volatility to generate a follow-up move large enough to generate a profit.
- Must have a favorable likelihood of success.
- Must hedge another position, thereby reducing overall risk.

When studying these events, there should be a direct relationship between the size of the reaction and the type of pattern that follows. For example, a small reaction to news may be followed by a steady continuation of the direction of the price shock. Therefore, if Initial Claims or total unemployment jumps by 0.2% in a month, the market expects a reaction by the government to stimulate job growth by lowering interest rates. The same initial reaction would occur if unemployment jumped by 1 full percent, but the number would be so unexpectedly large that it may be considered an error, in which case it is not clear to what extent the government would respond. The market may overreact to a large shock but underreact to a small one. The best way to discover this is to study the reaction to events.

We could reason that, if trend-following is profitable for a market, and all moves are the results of price shocks of different sizes, then the normal price shock will occur in the direction of the trend.

Fundamental to understanding price shocks that occur as a reaction to economic reports or corporate earnings, is that the size of the shock is based on the difference between the *expectations* and the *actual* reported data. The best assessment of
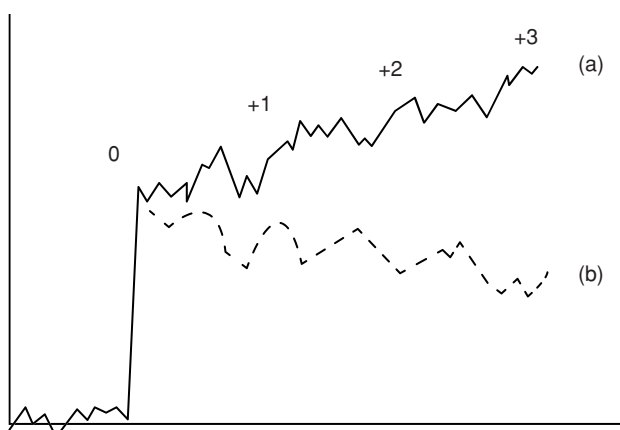
FIGURE 14.1   Price reaction to unexpected news, including delayed response. The upwards price jump on the event day (0) may be followed by (a) a continuation move up or (b) a reverse move down over the next 3 days.

expectations is the market price just ahead of an economic report. The market always discounts what it believes is its best guess at what the report will say; therefore, if bond prices rise in advance of an important unemployment report, we can say that the market expects unemployment to increase. If the futures price of short-term rates have moved up by the equivalent of ½% drop in yield in anticipation of a very bad report, and the report comes out neutral, then prices will drop sharply to offset the incorrect expectation. When studying market reactions from historic records of economic data, you must use price to represent expectations. A study of how much prices have moved between releases of the same economic report can tell you whether the current price fully discounts the anticipated announcement.

Measuring an Event

The most practical way of measuring the surprise effect of an event is to compare the volatility on the day of the shock to the average volatility. This can be done by comparing the volatility of the current day with the average volatility of the recent past, using the *true range* calculation:

$$Volatility\ ratio_t = \frac{True\ range_t}{Average\ true\ range_{\text{previous } n\text{-days, lagged } m\text{-days}}}$$

where the average true range is calculated over n days but the value used is the average lagged m days. Typically, $n = 60$ and $m = 30$. The use of 60 days for an average will reduce the effect of any other price shocks that occur within that period, and the lagging avoids missing two shocks that follow one another. When the ratio is greater than 3.0 we are likely to have identified a day in which a news event caused a significant price shock. The larger the ratio, the greater the surprise to the market. Note that the true range is used because the shock can occur overnight, causing an opening gap, in which case the best measurement of volatility includes the price move relative to the previous close. To get the best results, use only *pit session* data.

Volatility Pattern

During the declining stock market beginning in 2000 and subsequent rally, the key economic reports watched by traders were weekly Initial Claims and the monthly Unemployment. One study of volatility shows that price movement in the Treasury futures is greatest on the day of the report and lowest on the day after the report. There are cyclic variations in volatility that seem to occur in a weekly pattern between the monthly releases of these reports.

Trading the Event Lag

Once you have studied the way prices move following an event day or a price shock, you can create trading rules. For example, you find that a monthly change in the Consumer

Price Index of between 0 and 0.2% has no particular reaction, provided market expectations were similar to the data released. A change between +0.2 and +0.4 or –0.1 and –0.4 should provide enough surprise for a noticeable market reaction, when accompanied by lower or higher expectations. Data greater than a change of ±0.4 should cause a price shock regardless of the anticipation. In addition, we might find that changes over +0.4 tended to react in the opposite direction to the normal government action because of overanticipation or disbelief.

Based on the results of studying these patterns, Warwick established the following trading rules:

1. Buy on the close if the market closes in the upper 20% of the trading range on an event day, and if prices have shown an historic continuation pattern.
2. Hold for a predetermined number of days, based on the lag pattern of the market.
3. Use a stop-loss to limit risk.

In an older study applied to the corn market, Arnold Larson[6] found that 81% of the price changes occurred on the event day. There was a typical price reversal of 8% over the next 4 days, and a net change in price of 27% in the original direction over 45 days. An update of this study, shown at the end of this section, confirms that human behavior is not easily changed.

Results of Event Studies

In his book *Event Trading,* Warwick showed the test results of some of the major markets, allowing us to compare the effects of the systematic reaction of U.S. 30-year bonds with that of the S&P 500 and other financial markets. Unexpected news that causes bond prices to rally sharply (causing yields to decline) would result in a rise in the stock market, unless the news was particularly bad for the long-term economy of the country. Therefore, a drop in the consumption of durable goods should be followed by lower interest rates and a slightly lagged upwards response in the stock market. This pattern is shown in Table 14.1 where the bond and S&P results are side by side.

In general, we expect a price shock to impact interest rates first, then currencies, and finally equities markets. Typically, the interest rates react in a way to offset the economic effects and stabilize the equity market; this preempts the action that would be expected from the Central Bank. The foreign exchange (FX) markets react positively to raises in rates, and negatively to decreases, tempered by the reason for the rate change. The exception is when the news is much more extreme than usual and targets the stock market first. Investors, because of a lack of confidence, will shift their funds from equities to interest rates in what is called a *flight to safety*. This will cause a decline in the stock

---

[6]Arnold Larson, "Measurement of a Random Process in Futures Pricing," *Food Research Institute Studies* 1, no. 3 (November 1960), referenced in Warwick, *Event Trading* (1996).

market and a corresponding drop in yields, although Warwick normally expects the two markets to move in the same direction. Lower yields (higher bond prices) should lower debit and facilitate business, thereby raising stock prices. However, investors are cautioned that the time lag in the reaction of Treasuries and equity index markets may vary considerably due to the state of the economy and current interpretation of economic policy.

Table 14.1 shows the way the markets move after both the interest rates and stock market initially react with an upwards price shock. When both markets continue in the same direction, such as following retail sales, GDP, and employment data, then the full impact of the data takes longer to be assessed by the market, or there is expectation of some continued response by the government. The PPI and National Association of Purchasing Managers (NAPM) reports do not show this consistency, indicating that the market takes this information differently. When the pattern of confidence is low, the market may do a good job of reaching the best price level in immediate response to the news; therefore there is no consistency in the price move that follows.

Table 14.2 shows the side-by-side test results of reactions that cause a drop in bond prices and a corresponding drop in the S&P 500. The frequency of low-confidence results may represent the conflict between the normal upwards bias of the stock market or the unusually bullish trend during the test period. The GDP results that show the most confidence indicate that the stock market moves higher even when the interest rates rise in reaction to positive GDP data. Balance of Trade data, which is not always the primary focus of the market, has a very consistent pattern, but is not clearly related to the interest rate move as are most of the other statistics.

Although this work concentrates on the short-term reaction to unexpected news, the long-term trend should not be overlooked. Economic data can exhibit trends over long time periods, and the response by the Fed is usually moderate but consistent. If the CPI or PPI show early signs of inflation, then the monetary authority will push rates up slightly; if this doesn't work, as seen in the next series of economic reports, they will

TABLE 14.1   Results of Upwards Breakout of U.S. Bonds (left) and S&P 500 (right) for the Period 1989–1994

	U.S. Bonds			S&P 500		
	Holding Period	Buy or Sell	Confidence Level	Holding Period	Buy or Sell	Confidence Level
Durable goods	0–1	B	97	0–4	B	> 99
Retail sales	0–1	B	90	0–2	B	96
Retail sales	1–5	S	95	2–4	S	96
CPI	Low confidence			0–3	B	92
PPI	0–5	S	95	0–1	B	97
NAPM	0–2	B	91	0–4	S	97
Industrial production	0–1	S	94	Low confidence		
GDP	0–3	B	90	0–4	B	96
Employment	0–2	B	91	0–1	B	93

TABLE 14.2 Results of Downwards Breakout of U.S. Bonds (left) and S&P 500 (right) for the Period 1989–1994

	U.S. Bonds			S&P 500		
	Holding Period	Buy or Sell	Confidence Level	Holding Period	Buy or Sell	Confidence Level
Durable goods		Low confidence		0–2	B	92
Retail sales		Low confidence		0–4	B	94
Retail sales		Low confidence			Low confidence	
CPI	0–4	S	97		Low confidence	
PPI		Low confidence			Low confidence	
NAPM		Low confidence			Low confidence	
Industrial Production	0–4	S	93		Low confidence	
GDP	0–3	S	91	0–3	B	>99
GDP	3–4	B	>99	0–3	B	>99
Employment	0–1	S	93		Low confidence	
Trade balance	0–3	S	93	0-3	S	>99

move rates slightly higher again. This pattern creates an underlying trend that can be used to filter the direction of trades and add confidence to the results.

Readers are cautioned that the Treasury-stock index relationship can change for long periods due to the state of the economy. While the underlying economic principles will eventually surface, it is always important to study the market yourself and verify the patterns that you will use for trading. The process will build understanding and confidence in the method.

Reaction to Unemployment Reports

During the recent economic downturn, unemployment has been the focus of the recovery, and tracking those people filing unemployment claims has been the key measure in the past. Using Logical Information Machines (LIM), it was found that when the market underestimated its expectations of the monthly unemployment figure (issued on the first Friday of the month) by at least 0.2%, there was an average return of 4.5% in the S&P 500 during the following 27 days from the close of the report day. Similarly, U.S. 30-year Treasury bonds showed an average return on a long position of 3.2% during the period from 3 to 28 days following the report.[7] The interpretation is that the economy is worse than expected and will continue that way.

Trading the Reaction of Treasuries to Economic Reports

Treasury prices are often the market most affected by economic reports. A good report implies that interest rates will rise, and a poor report will cause them to fall; the reaction presumes that the government will use interest rates to counterbalance economic

---

[7]Gibbons Burke, "Event-Based Analysis," *Futures* (April 1995).

growth, above or below its target level. As discussed in the previous sections, futures markets reflect the expectation of these reports. In a study by Ruggiero[8] that covered 175 report days, Treasuries were seen to overreact to the numbers released in the Producer Price Index (PPI), Unemployment, and Retail Sales on the day of the report, plus the days of Treasury auctions.

To profit from the reaction of Treasuries, Ruggiero used a 1-day raw stochastic, *(close-low)/(high-low)*. He then bought Treasuries when the value was below 0.50 and sold when it was above 0.50 on the next open following the report day. Positions were held for 3 days. Based on the results of bonds, the reaction to the PPI was best, retail sales next, and unemployment last, although still profitable. We can conclude that traders overreact to all three reports, but higher volatility will yield better returns.

Raschke Trades the News

In her popular book with Laurence Connors, Linda Raschke[9] trades only the reaction to the morning economic reports. Using 30-year U.S. bonds as an example, if a report such as the Producer Price Index or Gross Domestic Product causes bond prices to jump more than $\frac{4}{32}$ above the high of the previous day, then

- Place a sell stop $\frac{3}{32}$ below the high of the previous day to enter a short sale on a reversal.
- Once the new trade is entered, place a stop-loss $\frac{1}{32}$ above the high of the current day.
- Move that stop down as soon as possible to a break-even point.

Raschke takes a businesslike approach to trading, keeping risk as low as possible and the chances of success as high as possible, even during the unusually high volatility that follows the release of morning reports. The same method is suggested for currencies, but there is no reason why it could not be used for any market that is directly affected by the economic numbers, such as stock indexes. European interest rate and index markets might be particularly interesting because they also react to U.S. reports but then must adjust to their own economic situation.

Finding Recent Patterns

We can write a program to find the patterns that follow any price shock. As discussed earlier, the average true range will allow us to find relatively large shocks. Start with the 60-day average true range, then calculate the ratio of today's true range divided by the average true range lagged 30 days (lag = 30). The subscript *ndays* is needed because we inspect the ratio *ndays* ago and look to see how price moved from *ndays* + 1 to today.

[8]Murray A. Ruggiero Jr., "Exploiting Report Day Tendencies in Treasuries," *Futures* (February 2001).
[9]Laurence A. Connors and Linda Bradford Raschke, *Street Smarts* (Malibu, CA: M. Gordon Publishing, 1995).

```
                    ATR = AvgTrueRange(60);
              if ATR[lag] = 0 then ratio = 0
           else ratio = truerange[ndays] / ATR[lag];
```

Whenever the *ratio* > *T*, where *T* is a threshold value of, for example, 2.0, we print the ratio, the price change on the day of the price shock, then the cumulative price change starting on the close of the day of the price shock (shown as *Start* on the *x*-axis). The cumulative price changes show what have happened if a long or short position had been taken on the day of the price shock and held for 10 days. The program for this, *TSM Price Shock Reactions*, can be found on the Companion Website.

Results are shown for three different markets over the past 10 years, ending April 2011. The threshold level for the ratio was 2.0, indicating a price shock in which today's volatility was at least twice the average volatility, all based on the closing prices. Figure 14.2a shows that the largest positive and negative price shocks during the last 10 years occurred in 2008 and 2009. The starting point on the chart shows the closing gap on the day of the shock, $5\frac{2}{32}$ higher and $3\frac{15}{32}$ lower. Positions are taken at the close of the day of the shock, and the accumulated net price change shown on days 2 through 10. After 10 days the downward price shock ends more than 1 point higher while the upwards price shock shows only a small gain. At its best point, we could have gained $1\frac{16}{32}$ in 4 days by selling the upwards price shock on the close. We could have also gained the same amount in 5 days by buying the downward shock on the close. It was mentioned earlier that volatility drops off quickly, and this chart shows that the accumulated reaction to the largest price shocks is moderate. Both shocks end higher after 10 days, showing that the long-term trend surfaces once the impact of the shock fades.
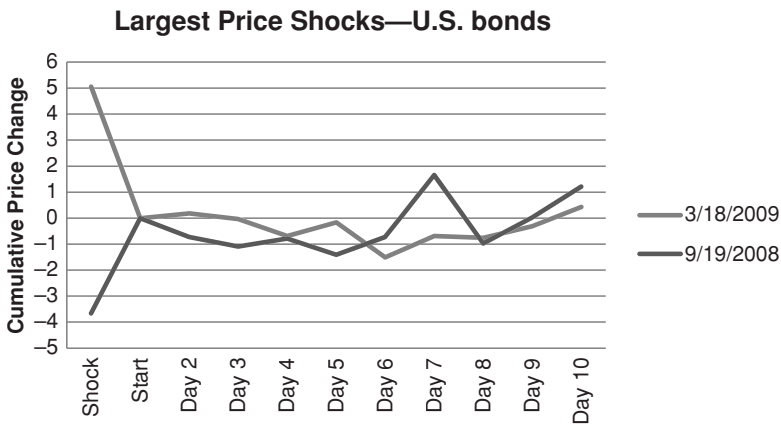
FIGURE 14.2a U.S. 30-year bond price shocks and cumulative reaction for 10 years ending April 2011. (a) The two largest upward and downward shocks, (b) the average of all shocks with a closing price move of more than 1.5 points (48/32nds).

**Average Net Change in U.S. Bonds
following a 1.5 pt Move**

(b)
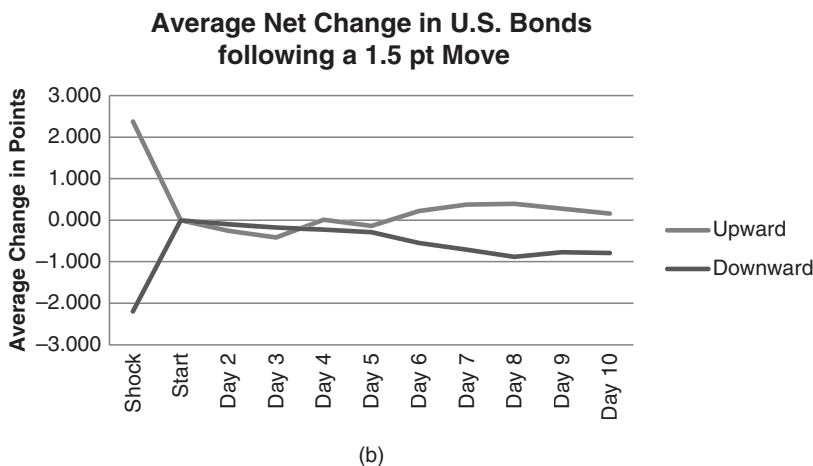
FIGURE 14.2b   Average shock > 48/32nds and subsequent cumulative
price changes.

Looking at the largest shocks is not representative of the returns from trading. Figure 14.2b shows the average of all price shocks greater than $1^{16}/_{32}$. It shows a steady move back in the direction of the shock. The downward shock is the most profitable, with an average return of nearly 1 full point by selling on the close of the day of the shock. The average volatility is very low. This is very similar to the case shown in Figure 14.1 where price continue slowly in the direction of the shock for a few days afterward.

The underlying direction of the market should be considered when evaluating results. In the case of Warwick's study, from 1989 to 1994, there is a shorter, volatile sideways period followed by a significant bull move. You would expect greater confidence buying downward price shocks. In the pattern found in Figure 14.2b we would also expect either upwards or downwards shocks to be followed by an upwards move, but that doesn't happen. Instead, they are quite symmetric, a sign that the underlying direction did not bias the results. A chart of 30-year bond futures is shown in Figure 14.3.

**FIGURE 14.3**   U.S. 30-year bond futures, 1986 to April 2011.
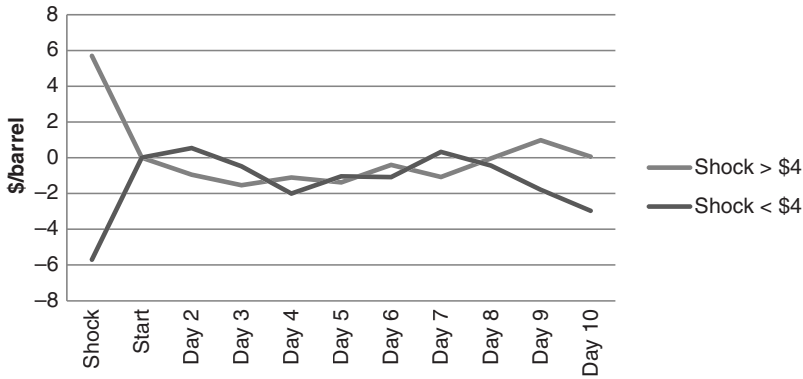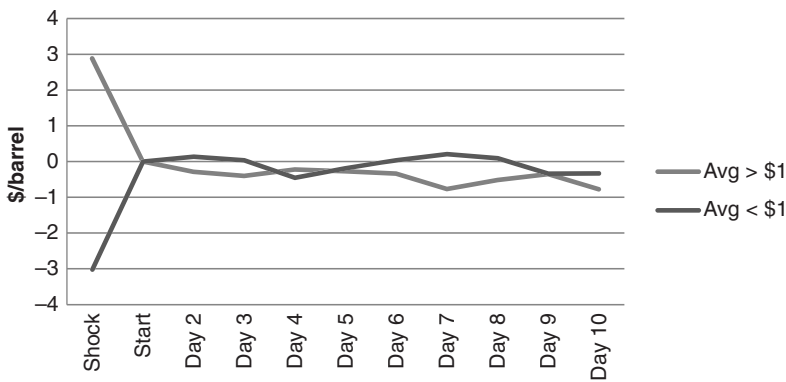
FIGURE 14.4 Results of price shocks in crude oil futures for 10 years.
(a) Shocks greater than $4/barrel, and (b) shocks greater than $1/barrel.

The results of crude oil tests are shown in Figure 14.4. The price shocks in the crude market are different from bonds because they most often come from geopolitical events rather than weekly API or economic reports. For large shocks, where the market closed higher or lower by at least $4/barrel, averaging nearly $6/barrel, the biggest opportunities were selling. Both upward and downward shocks ended nearly $2 lower after 3 and 4 days and although the downward shocks ended much lower after 10 days, we should assume that it was reverting to the trend. The memory of a price shock should be less than 10 days.

To see if the large price shocks are different from the more frequent, smaller shocks, Figure 14.4b shows the average of all shocks greater than $1. The results are muted compared to the shocks greater than $4 and the upside price shocks have a modest downside reaction while the downside price shocks offer little opportunity.

The S&P is of interest to nearly everyone, and Figure 14.5 shows that upward price shocks greater than 30 points are followed by steady substantial declines, netting about 40 big points, or $2000 per contract. Downward price shocks fail to produce either a
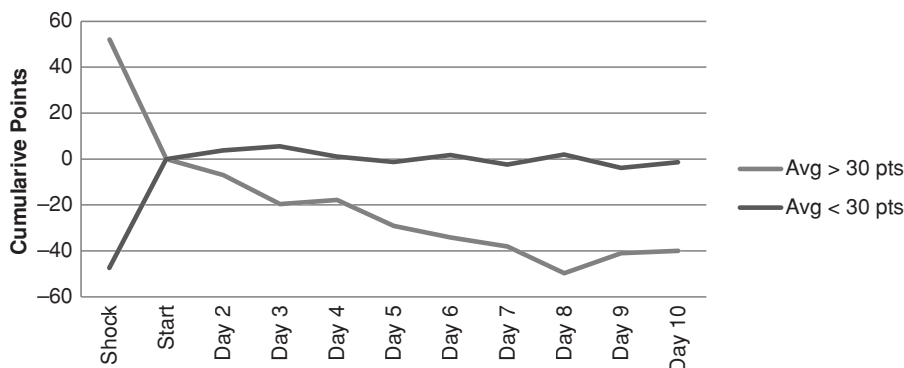
FIGURE 14.5 S&P price shocks greater than 30 points.

profit or a loss, indicating that once a large drop occurs, prices sit at that level for a number of days. It should be noted that a shock of 30 points is more likely to occur at higher price levels, which represent a smaller percentage. During the past 10 years, the price of the S&P index has fluctuated higher and lower, essentially ending where it began. The period of highest frequency of price shocks, both up and down, occurred in the fall of 2008, during the subprime mortgage crisis. Although the pattern was more prevalent years ago, there should still be a bias for price shocks to the downside. Most traders/investors prefer to be long the equity markets, causing a vacuum underneath when a large number of them want to exit at the same time.

The details for these tests are available on the Companion Website in spreadsheet files labeled *TSM Price Shock Reactions*, followed by the market code.

Conditional Analysis of Shocks

So far, the analysis of shocks has only considered the direction of the closing price, above or below the previous close by some large threshold value. Even with daily data, there are patterns that can refine this study and make it more practical. What if on the day of the large move, the price closed near the high or near the low? If the shock was upwards, and prices closed near their highs, then we might expect a continuation of that move into the next day. If an upwards shock closed near the lows of the day we might expect a reversal.

Using U.S. 30-year bonds as an example because it has frequent shocks of varying size, a test was run using a 1.5 threshold for the price shock (prices close 1.5 times higher or lower than the average true range), and the close was in the highest or lowest 20% of the daily trading range. Figure 14.6 shows that the results for the 6 days following the shock are much different from the study that ignores the relative position of the close. For upwards shocks, part a, a lower close has the distinct tendency to continue lower, although an average of only 17 basis points, while a higher close fails to follow through. For downward shocks, part b, the upward reaction is much stronger when prices close near the highs of the day. During the two days after the shock, prices gain 50 basis points
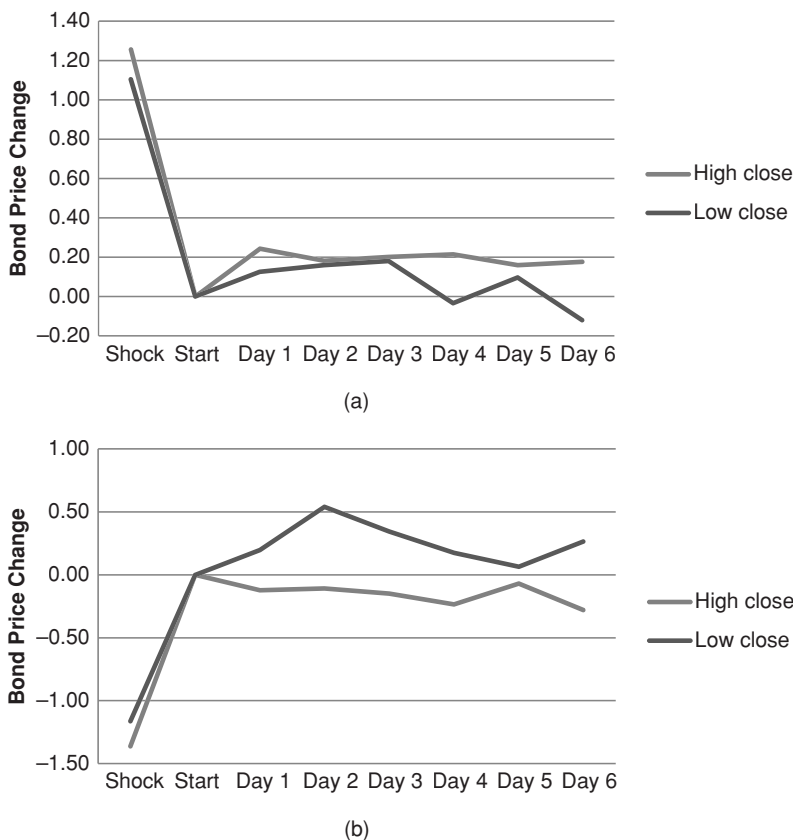
FIGURE 14.6 Results of price shocks of 1.5 with prices closing in the highest or lowest 20% of the daily trading range. (a) Upwards price shocks. (b) Downwards price shocks.

before settling back to the original level of the price shock. Keeping these reactions in context, bonds have been in an uptrend for 25 years, so that an upwards move following a downwards price shock could simply be a reversion to the trend.

COMMITMENT OF TRADERS REPORT

Drawing from more than 20 years of experience in analyzing the CFTC's *Commitment of Traders Report* (COT), William L. Jiler described an approach to identifying a major trend in his *1985 CRB Commodity Yearbook*.[10] With his usual thoroughness and clarity, Jiler presents material which had previously been unsuccessfully interpreted.

---

[10] William L. Jiler, "Analysis of the CFTC Commitments of Traders Reports Can Help You Forecast Futures Prices," *1985 CRB Commodity Year Book* (Jersey City, NJ: Commodity Research Bureau, 1985).

Originally released on the 11th day of each month, the publication of the COT report changed to twice each month in 1990, biweekly in 1992, and has been released every Friday since January 2000. The content of the report has remained essentially the same; however, in 2010 the CFTC made a new attempt to classify the traders more accurately. The *short form COT Report* summarizes the open interest (open positions) of *reporting* and *nonreporting* traders as of the close of trading on the Tuesday of the reporting week. *Reported positions* are those exceeding a minimum level determined for each futures market (e.g., 750,000 bushels for corn, 500,000 for soybeans, 500 contracts for 30-year Treasury bonds, and 300 contracts for the S&P 500). These levels are based on the total activity of the market and may change over time. As of January 2000, those traders with reportable positions must aggregate all smaller positions held in the same market into one number. Previously, only the individual orders in excess of the reporting limit needed to be identified. Additional information is available that subdivides the open interest into positions held by hedgers, who are commercial users of the commodity, and speculators. By subtracting the reported positions from the total open interest, the total positions of small hedgers and speculators can be found.

The short form, shown in Figure 14.7, gives a brief breakdown of speculative and commercial positions. In the long form, Figure 14.8, grain positions are further divided into *old crop* and *other* positions, including new crop and intercrop spreads. At the bottom are the concentration of positions held by the largest four and eight traders.

In order to analyze the shifts in position, Jiler compiled tables of *normal patterns*, similar to a seasonal study (see Figure 14.9). When the open interest of one group is significantly greater than their normal holdings, they express a definite

WHEAT – CHICAGO BOARD OF TRADE

FUTURES-ONLY POSITIONS AS OF 12/12/06

| NONCOMMERCIAL			COMMERCIAL		TOTAL		NONREPORTABLE POSITIONS	
LONG	SHORT	SPREADS	LONG	SHORT	LONG	SHORT	LONG	SHORT
(CONTRACTS OF 5,000 BUSHELS)		OPEN INTEREST: 417,081						
COMMITMENTS								
73,598	56,045	69,448	237,539	232,901	380,585	358,394	36,496	58,687
CHANGES FROM 05/25/2004		CHANGE IN OPEN INTEREST: –7,043						
–10,463	–1,186	126	3,462	–6,610	–6,875	–7,670	–168	627
PERCENT OF OPEN INTEREST FOR EACH CATEGORY OF TRADERS								
17.6	13.4	16.7	57.0	55.8	91.2	85.9	8.8	14.1
NUMBER OF TRADERS IN EACH CATEGORY (TOTAL TRADERS: 317)								
102	89	92	67	96	233	256		

FIGURE 14.7   Commitment of Traders Report (short form) from the CFTC.gov website

WHEAT – CHICAGO BOARD OF TRADE
Commitments of Traders – Futures Only, May 10, 2011

	Total Open Interest	Reportable Positions							Nonreportable Positions	
		Non-Commercial			Commercial		Total			
		Long	Short	Spreading	Long	Short	Long	Short	Long	Short
	(CONTRACTS OF 5,000 BUSHELS)									
All	455,903	100,101	89,939	65,413	253,366	243,455	418,880	398,807	37,023	57,096
Old	415,581	100,159	95,463	46,454	237,460	227,124	384,073	369,041	31,508	46,540
Other	40,322	11,869	6,403	7,032	15,906	16,331	34,807	29,766	5,515	10,556
	Changes in Commitments from May 3, 2011									
	−4,544	−4,092	−1,627	2,347	−3,085	−8,434	−4,830	−7,714	286	3,170
	Percent of Open Interest Represented by Each Category of Trader									
All	100.0	22.0	19.7	14.3	55.6	53.4	91.9	87.5	8.1	12.5
Old	100.0	24.1	23.0	11.2	57.1	54.7	92.4	88.8	7.6	11.2
Other	100.0	29.4	15.9	17.4	39.4	40.5	86.3	73.8	13.7	26.2
	Number of Traders in Each Category									
All	421	123	132	153	84	132	311	337		
Old	418	121	141	138	79	127	294	326		
Other	171	37	31	19	33	86	84	129		

	Percent of Open Interest Held by the Indicated Number of the Largest Traders							
	By Gross Position				By Net Position			
	4 or Less Traders		8 or Less Traders		4 or Less Traders		8 or Less Traders	
	Long	Short	Long	Short	Long	Short	Long	Short
All	22.3	18.6	37.4	25.5	20.0	13.5	34.1	19.0
Old	23.5	20.2	39.5	27.7	22.1	15.8	36.7	22.3
Other	32.5	23.5	48.5	31.8	25.1	14.8	38.5	20.5

FIGURE 14.8 Commitment of Traders Report (long form), May 10, 2011.

opinion on the direction of the market. By tracking these changes for many years and observing the corresponding price changes, Jiler concludes that the large traders have the best forecasting record, with the large hedger better than the large speculator. The small traders were notably worse. Guidelines, referring to agricultural products, were stated as:

> *The most bullish configuration would show large hedgers heavily net long more than normal, large speculators clearly net long, small traders heavily net short more than seasonal. The shades of bullishness are varied all the way to the most bearish configuration which would have these groups in opposite positions—large hedgers heavily net short, etc. There are two caution flags when analyzing deviations from normal. Be wary of positions that are more than 40% from their long-term average and disregard deviations of less than 5%.*
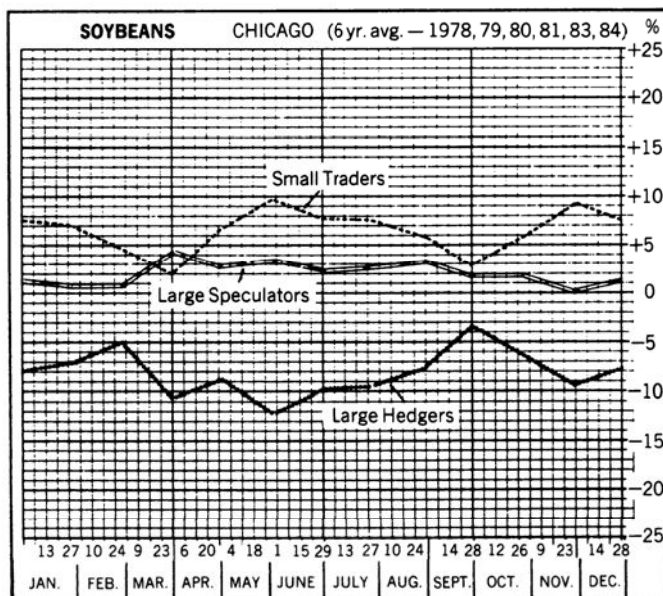
**FIGURE 14.9** Jiler's normal trader positions.

This result was confirmed by Curtis Arnold,[11] who compared positions of large and small speculators with commercials for 1-year periods spanning 1983–1984. Arnold shows in Figure 14.10 that the position of commercials and small speculators tend to be opposite, with the commercials positioned to profit from the subsequent price move.

Over a wider group of 36 markets from 1983 to 1989, a study conducted by *Bullish Review*[12] showed that a large weighting of long or short positions held by commercial hedgers correctly forecast significant market moves 67% of the time. In retrospect, it seems likely that a large commercial trader, bank, investment house, or other institution would take a longer-term view of the market, setting positions to take advantage (or avoid the risk) of a potential price move. Because of their size, institutions are not likely to change their positions on minor price variations.

Analysts should be aware that a qualified *hedger*, one that is registered with the CFTC in order to set large positions due to their ongoing risk of holding physicals, may also take large speculative positions under the same umbrella. Given this ambiguity, the new COT classifications will make it clearer where the "smart money" is positioned. The old report combined commercials and swap dealers; however, swap dealers can side with the speculator or commercial given market conditions. By separating them, the report now shows "true commercials."

---

[11]Curtis Arnold, "Tracking 'Big Money' May Tip Off Trend Changes," *Futures* (February 1985).

[12]Steve Briese, "Tracking the Big Foot," *Futures* (March 1994).
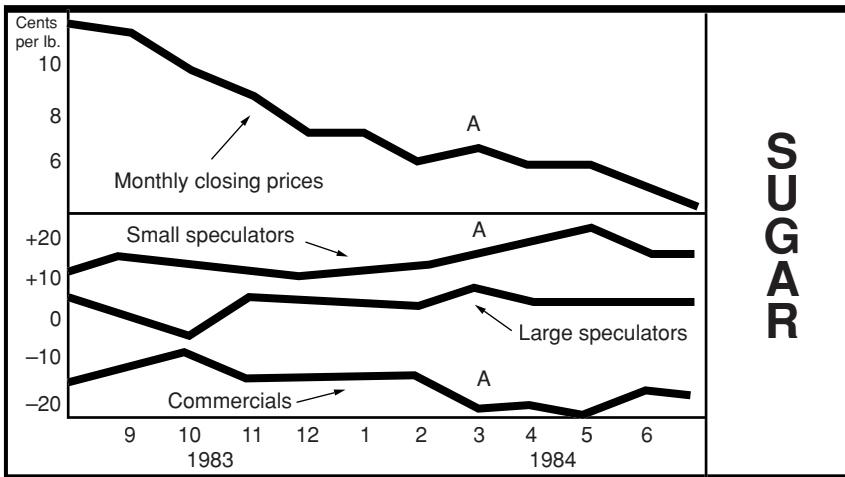
FIGURE 14.10 Arnold's trader position study.

Creating an Oscillator for the COT Numbers

Briese uses a *COT Index*[13] as an oscillator to plot the results of the COT report.

$$COT\ Index_t = 100 \times \frac{NL_t - Lowest(NL, n)}{Highest(NL, n) - Lowest(NL, n)}$$

where NL = the net long positions of a given group of reporting traders
        n = the look back period, ranging from 1.5 to 4 years

This should be recognized as the formula for a stochastic, giving the bullish percentage.

In order to see the cycle of the three reporting groups, Ruggiero[14] plotted the stochastic of each group using a calculation period of 1 year. The positions of commercial and those of small-lot traders are completely out of phase. The stochastic indicator shows that the commercial traders reach an extreme level of bullishness (high stochastic values) just ahead of an upwards price move, and a low stochastic value just before prices drop. This implies that they are usually holding the right position. Ruggiero interprets the other values to mean that funds trading these markets become bullish or bearish when the commercials are beginning to exit their positions, about halfway through the move. Small-lot traders tend to become bullish, with a high stochastic reading, as prices peak.[15]

---

[13]In Murray A. Ruggiero, Jr., *Cybernetic Trading Strategies* (New York: John Wiley & Sons, 1997). Also see Stephen Briese, *Commitment of Traders Bible* (Hoboken, NJ: John Wiley & Sons, 2008).

[14]Murray A. Ruggiero Jr., "Seeking a Commitment," *Futures* (April 2002).

[15]The same conclusion is reached by Jon D. Andersen in "Analyze Net Positions," *Technical Analysis of Stocks & Commodities* (July 2002), although Andersen plotted the actual COT values—rather than convert to an indicator—and observed the diverging patterns.

While this study is consistent with previous works, the use of a stochastic makes a strong visual impact, showing how the bullishness and bearishness of these different groups are remarkably out of phase. The best short sale signals occur at just these times, when the positions of the commercials are strongest, well above 80, and the two other groups are weakest, under 20. Buying opportunities are just the opposite. Extreme differences precede change in the positions of each group and the market price. The holding time for this trade corresponds to the time it takes for the stochastic to switch from high to low, and low to high, for the different groups. This may take 3 to 6 months, depending on the market.

Trading Signals

An additional trading refinement was applied to the reported positions for the S&P 500. When the stochastic for the commercials crosses above 80 and the S&P closes lower than it did 4 weeks ago, we get a buy signal. The position is closed out when the stochastic touches 40, which is just before the neutral level of 50. It is not safe to assume that stochastic values will move from overbought to oversold; it is much better to only expect them to return to neutral.

Ruggiero also used the COT Index for both commercials and small traders to create a basic timing model:

If COT Index Commercials[lag] > *Ctrigger* and COT Index Small < *Strigger* then *buy*

If COT Index Commercials[lag] < *Ctrigger* and COT Index Small > *Strigger* then *sell*[16]

The results are bullish when the COT Index for commercials is high and the COT Index for small traders is low. The results are bearish when the COT Index for commercials is low and the COT Index for small traders is high. The COT Index for commercials is lagged by one or more weeks because the actions of commercials are interpreted as leading the market. *Ctrigger* and *Strigger* are levels that are determined to be significant turning points. In Ruggiero's tests, this method was highly successful for Treasury bonds using *Ctrigger* = 30, *Strigger* = 50, and *lag* = 1 (one week).

Capturing the Cycle in the COT

In another more extensive strategy, Ruggiero created a method that anticipated price direction based on extreme bullishness and bearishness in the COT report; however, he chose to only take long positions.[17] The pattern being exploited is called the *markup phase*, which occurs after the commercials have stopped buying and start selling the rallies. At the same time, funds enter the market as it accelerates. The public enters when prices are very bullish and commercials are turning bearish.

---

[16] In Ruggiero, *Cybernetic Trading Strategies*, this was written as a *buy* signal; however, it seems that it should have been a *sell* signal.

[17] Murray Ruggiero, "Trading off the Big Boys," *Futures* (January 1998).

The purpose of this strategy is to enter the market at the time of upwards acceleration and exit after at least n days and after the acceleration has stopped. To implement this, Ruggiero finds the correlation of price and volume and converts that to an oscillator rescaled to the range +1 to –1. He finds the Herrick Payoff Index (HPI) using a conversion factor of 100 and a smoothing constant of 0.10 (see Chapter 12), which combines price, volume, and open interest. A *markup* occurs when the correlation oscillator > 0.3 and HPI > 0 but not the highest value of the recent period and the payoff has not retraced by some threshold percentage.

Once the markup has been determined, another oscillator is formed from the net commercial position, longs minus shorts, ranging from 100 to 0. If the oscillator > 85 and a markup occurred within the past 3 days, a long position is entered. Trades are held for at least n days and until there is no markup. An upgraded version of Ruggiero's program for this strategy is *TSM COT Markup*, available on the Companion Website.

The numbers in the COT report can be turned into a *Sentiment Index*, as shown in the next section of this chapter.

OPINION AND CONTRARY OPINION

Market sentiment, the collective opinion of investors, is a driving force in the market, yet it is very difficult to measure and even harder to deliver those results in a timely fashion. For that reason, analysts often substitute a combination of volume, open interest, and price for true sentiment, hoping that the recorded actions of traders closely relate to what they are thinking. Opinion, however, weighs on the marketplace and governs future actions in the same way that high volume may not move prices today, but provides a platform for a potentially large price move. The *Consumer Confidence Index*, released monthly by the Conference Board and conducted by Nielson, is the most common example of opinion that moves the market.

Public opinion is also fast to change. A prolonged bull market in stocks may show a gradual increase in bullish sentiment; however, the collapse of a bank, an increase in rates by the central bank, a sharp downturn in the economy of another region, or a single sharp drop in the stock index could quickly change the public's opinion. In reality, sentiment indicators are most popular for trading in the direction opposite to the unified public opinion.

Contrary Opinion

The *contrarian* lies somewhere between the fundamentalist and the technician, basing actions on the behavior of crowds, in this case the market participants.[18] The contrarian

---

[18] For the most definitive works, see Humphrey Neill, *The Art of Contrary Thinking* (Caldwell, OH: Caxton Printers, 1960), who is credited with having first formulated the concept, and Earl R. Hadady, *Contrary Opinion* (Pasadena, CA: Hadady, 1983).

sees the end of a bull market occurring when everyone is bullish. Once all long positions have been set, there is diminishing influence by the bulls; moreover, the contrarian believes that opportunities always lie in the reverse direction from crowd thinking.

Contrary opinion alone is not meant to signal a new entry into a position; it only identifies situations that qualify. It lacks the timing. It is more of a filter than a trading system, a means of avoiding risk and finding an opportunity. Consider the patterns that appear in every prolonged bull or bear move. First, there is a place where the direction is generally accepted as the major trend. After that, traders wait for a pullback to enter in the direction of the trend at a more favorable price. These price *corrections* within a trend become smaller or even disappear when everyone wants to *buy a lower open* or in the case of a bear market, *sell a higher open*, until finally there is a *blow-off* and a major reversal. The dramatic end to a prolonged move is generally credited to the entrance of the public; when the masses are unanimously convinced that prices are going higher, who else is left to buy?

The other important ingredient for a contrarian is that all the facts cannot be known. The widely accepted belief that prices will go higher must be based on presumptions; if the final figures were out and the reality known, the market would adjust to the proper level. This idea is older than *The Art of Contrary Thinking.* In 1930, Schabacker[19] advised,
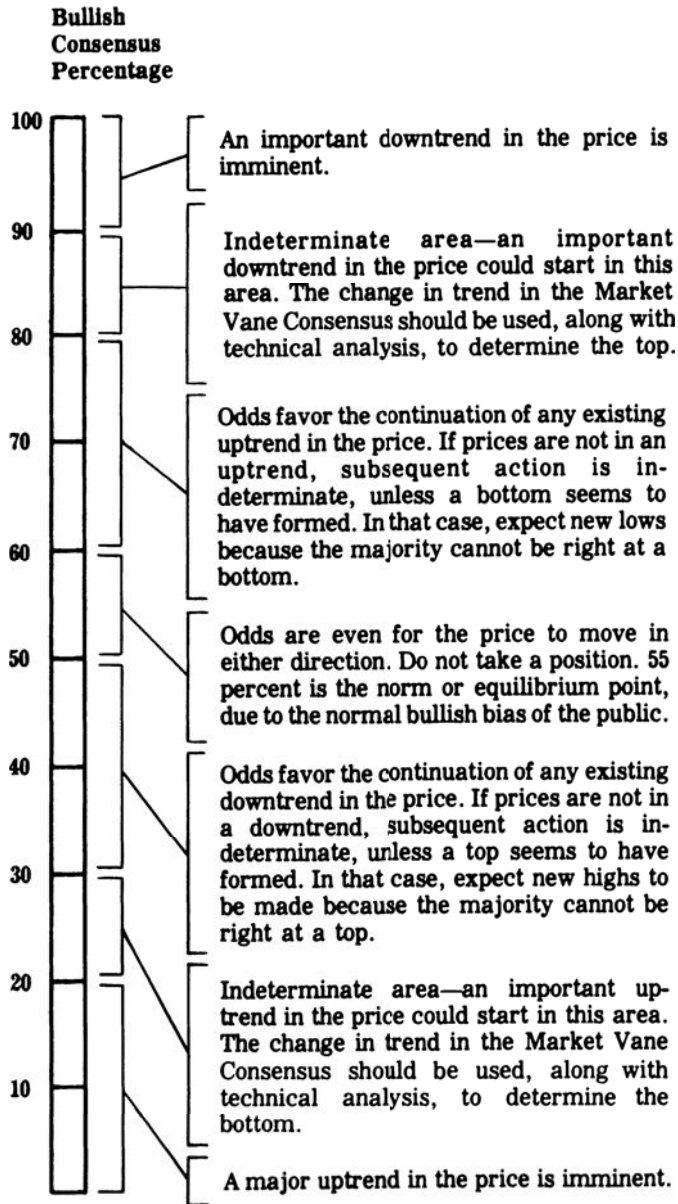
> *It is best to follow the crowd when fundamental considerations for the industry, for the stock and for the group are improving and indicate that there is really something to the move. But it is best to reverse the action of the crowd and desert them, or play the opposite side of the market, when their action is based more on rumors, on prospects, on paper talk, on public gossip, on results of pool publicity, etc., than upon a genuine raison d'être. Another clue is that when the resulting movement is sharper than usual and enthusiasm is higher than seems logical, it is also best not to follow the crowd.*

The practical application of the theory of contrary opinion is the *Bullish Consensus*[20] and the *Market Sentiment Index*, created from a poll of market letters prepared by the research departments of brokerage firms and professional advisors. In the *Bullish Consensus* (see Figure 14.11), these opinions are weighted according to the estimated circulation of these letters.

$$Bullish\ Consensus = \frac{Sum\ of\ (Each\ source \times Relative\ circulation \times Bullish\ opinion)}{Sum\ of\ (Each\ source \times Relative\ circulation \times 100)}$$

---

[19]R.W. Schabacker, *Stock Market Theory and Practice* (New York: B. C. Forbes, 1930), 517. This book was the definitive study of the market. In it he quotes Theodore Price as saying that speculation is "hazard plus intelligence" while gambling is "hazard without intelligence."

[20]The Bullish Consensus is a product of Sibbett-Hadady, Pasadena, CA; a Market Sentiment Index is published in Consensus, Kansas City, MO.

**Bullish
Consensus
Percentage**

100 — An important downtrend in the price is imminent.

90
80 — Indeterminate area—an important downtrend in the price could start in this area. The change in trend in the Market Vane Consensus should be used, along with technical analysis, to determine the top.

70
60 — Odds favor the continuation of any existing uptrend in the price. If prices are not in an uptrend, subsequent action is indeterminate, unless a bottom seems to have formed. In that case, expect new lows because the majority cannot be right at a bottom.

50 — Odds are even for the price to move in either direction. Do not take a position. 55 percent is the norm or equilibrium point, due to the normal bullish bias of the public.

40
30 — Odds favor the continuation of any existing downtrend in the price. If prices are not in a downtrend, subsequent action is indeterminate, unless a top seems to have formed. In that case, expect new highs to be made because the majority cannot be right at a top.

20
10 — Indeterminate area—an important uptrend in the price could start in this area. The change in trend in the Market Vane Consensus should be used, along with technical analysis, to determine the bottom.

— A major uptrend in the price is imminent.

FIGURE 14.11 Interpretation of the Bullish Consensus.

Source: R. Earl Hadady, *Contrary Opinion* (Pasadena, CA: Hadady, 1983).

The value of the bullish consensus ranges from 0 to 100%, indicating an increasingly bullish attitude. Because of the bullish tendency of the novice trader, and the long-term upwards bias of the stock market, the neutral consensus level is 55%. The normal range is considered from 30 to 80%, although each market must be individually evaluated.

Hadady also devised a simple mathematical way of displaying the bullishness of the market.[21] Using the formula below, he shows that when 80% of traders are bullish, then the average buyer will hold only ¼ the number of contracts as the average seller. This leads to a precipitous drop in prices once a decline begins.

$$N_t = \frac{20\%T \times N_s}{80\%T} = \frac{1}{4} \times N_s$$

where T = the number of traders in the market
         N_t = average number of contracts held by a single bullish trader
         N_s = average number of contracts held by a single bearish trader

The principles of contrary opinion do not require that trades only be entered in the direction opposite to the current price movement. Within the normal range, the contrarian will take a position in the direction of the trend. Frequently, the Bullish Consensus will begin increasing prior to the price turning higher, indicating that the attitude of the trader is becoming bullish. It is considered significant when the index changes 10% in a 2-week period. Once the Bullish Consensus reaches 90% during an upward move, or 20% during a bear move, the market is considered overbought or oversold, and the contrarian looks for a convenient point to exit from the current trade. Positions are not reversed until prices show that they are changing direction. This could be identified using a moving average. Remembering Schabacker's advice, the occasion of a general news release that moves the market further than would seem reasonable in the direction of the general opinion is an opportune moment to enter a contrary trade; specific news that fails to move prices would be a good indication of an exhausted trend when the consensus is overbought or oversold. A typical contrarian situation identified by Hadady is given in Figure 14.12.

R. Earl Hadady said, "The principle of contrary opinion, by definition, works 100% of the time. The problem is getting an accurate consensus."[22] Timeliness is the difficulty with this index; if 60 to 70 market letters are reviewed, read, and weighted to form an index, the results may be outdated before they can be used. The theory of contrary opinion also emphasized its use as a timing device for entering trades at an opportune moment and for filtering out the ambiguous trades; the theory is not readily applicable to exiting a position unless the reverse consensus occurs. Because a consensus does not have to switch uniformly from bullish to bearish, it is not always prudent to wait for an opposite confirmation before exiting a trade.

---

[21] R. Earl Hadady, "Contrary Opinion," *Technical Analysis of Stocks & Commodities* (August 1988).

[22] George Angell, "Thinking Contrarily," *Commodities* (November 1976).
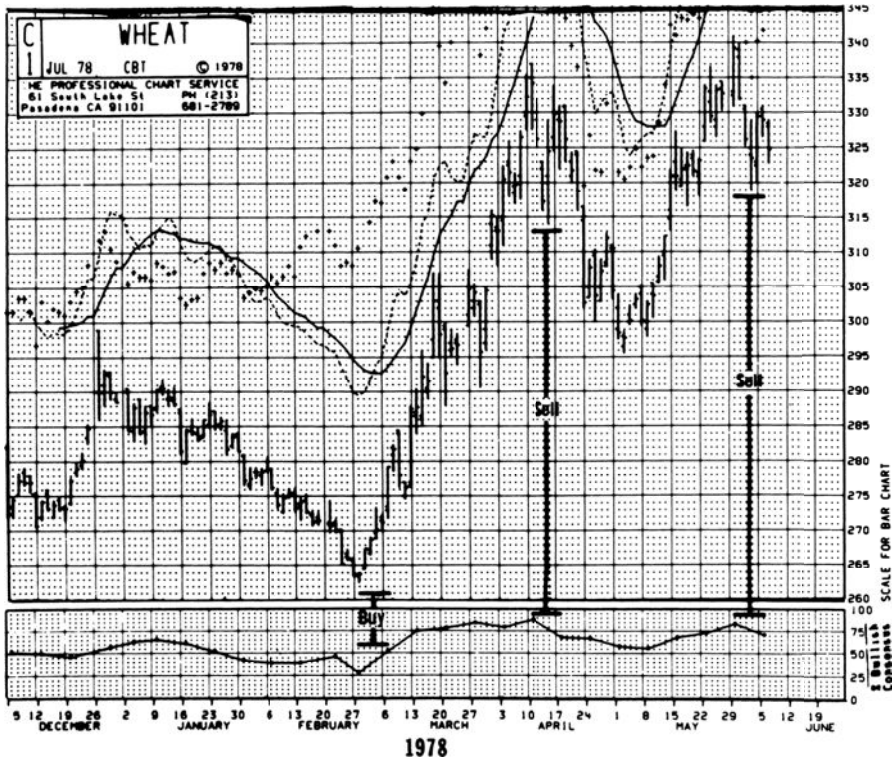
FIGURE 14.12 Typical contrarian situation—wheat, 1978.

Commitment of Traders Sentiment Index

The reported positions of traders, published in the CFTC's *Commitment of Traders Report*,[23] may be considered a recording of market opinion into categories. By combining the idea behind the stochastic calculation with a method originally developed by Curtis Arnold, you can create a *Sentiment Index*:

$$COT\ Sentiment\ Index = \frac{Current\ net - Minimum\ net}{Maximum\ net - Minimum\ net}$$

Where   *Net* = commercial net position (number of contracts) minus the total combined net position of large speculators and small traders

Maximum = greatest net difference that occurred during the comparison period

Minimum = smallest net difference that occurred during the comparison period

---

[23] Stephen E. Briese, "Commitment of Traders as a Sentiment Index," *Technical Analysis of Stocks & Commodities* (May 1990).

The intent of the Sentiment Index is to rank the current spread between the commercial and speculative positions within the context of the historic range, in a manner similar to a stochastic. There seems to be uniform agreement that the commercials are the group that determines the correct direction of prices. A shift in the position of the commercials should be closely watched.

A Lesson from Put-Call Ratios

The ratio of put option volume to call option volume, called the *put-call ratio*, is the major sentiment index for listed options. It too is used for its contrary value. Unlike the sentiment indexes derived from analyst opinions, this ratio is available daily. The interpretation of extreme levels of the put-call ratios, in particular for the major equity index options, points out a problem that may reflect on the proper use of all contrary indicators.[24]

Prior to 1986, the market was considered ready for an upturn when the total put volume exceeded 65% of the total call volume. Similarly, when the put volume fell to 35% of the call volume, it was a bearish indication. In the volatile markets of 1986 and 1987, these levels proved to be far too close, and as McMillan said, "Not surprisingly, the put-call ratios fell into some disfavor at that point." This could easily happen with a contrary indicator, or any indicator that rarely reaches its extreme values (even a relative strength index that has no maximum and minimum bounds). Because contrary opinion is a valuable addition to analysis, the use of these indicators now focuses on relative highs and lows. This can be accomplished by smoothing the ratio using a standard moving average or momentum indicator, a simple difference over n days. When the ratio moves over 65% *and turns down* it is time to sell. This approach gives up a timing edge but greatly reduces risk and increases reliability. After selling, if the momentum indicator reversed and made new highs, short positions should be exited or reversed.

Dogs of the Dow and the Small Dogs

Another recognizable contrarian method is the *Dogs of the Dow*, popularized by Michael O'Higgins in *Beating the Dow*.[25] The method shows that if, on January 1 of each year, you bought the 10 members of the Dow with the lowest stock dividend yields from 1973 through 1989, you would have beaten the average return of the Dow by 6.8% per annum. It assumes that good companies pull themselves up after a bad period. While not always profitable during the turn of the market between 2000 and 2002, it is reported to have done much better than the Dow Index.

For the small investor, O'Higgins proposed the *Small Dogs of the Dow*, also fondly called the *puppies.* Out of the 10 stocks selected as the Dogs of the Dow, buy the ones

---

[24]Lawrence G. McMillan, "Put-Call Ratios," *Technical Analysis of Stocks & Commodities* (October 1995).

[25]See Harry Domash, "Winning Investing.com," winninginvesting.com.

with the lowest stock value. In Robert Sheard's book, *The Unemotional Investor*, he presents the *Foolish Four*, a way of selecting 4 of the 10 Dogs.[26] As described on *winninginvestments.com*, Sheard sorts the 5 stocks of the Small Dogs, lowest price at the top, and places them alongside the 10 Dogs, lowest yields at the top. If the same stock appears at the top, it is eliminated because it is trading at a low price for a good reason. If there is no match, he takes the top 4 stocks from the Small Dogs.

Watching the Big Block Transactions

Big block transactions are those that are equal to or greater than 100,000 shares or have a transaction value of $1 million or more. *Corporate insiders* are the corporation's officers, directors, and beneficial owners of more than 10% as specifically named in the 1934 Securities and Exchange Act, Section 16. Corporate insiders must register their holdings with the SEC and must disclose any share transactions by filing a report within 30 days after the purchase or sale. Due to slowness in these filings, the Sarbanes-Oxley Act became effective in August 2003 to accelerate reporting (among other things), which now must be done within 2 days of the transaction. Insiders can also file a plan with the SEC, in advance, to buy or sell shares on specific dates or at specific prices. Not all of the shares traded by insiders are "timed."

Insider trading is not the same as *trading on inside information*, which is illegally making a trade based on nonpublic or privileged information. Insider selling is a common occurrence, as executives who have accumulated stock options, and venture capitalists who see the appreciation in share value, convert their holdings into cash. Some of these transactions are done at regular intervals and have no relationship to market direction, while others are at opportune moments and may offer insight into the future direction of prices.

The combination of big block trades by corporate insiders is a special case that was studied by Bjorgen and Leuthold,[27] who concluded in 1998 that

- Since 1983, when net selling, measured in dollars, reached historically high levels, the stock market performed poorly over the next 12 months. At the time this article was written, 1998, corporate insiders were selling at levels approaching historic highs.
- When net selling reaches historic lows, the market has performed significantly above average over the next 12 months.

The authors also noted that the number of net buy or sell transactions, regardless of total dollars, may also follow the same relationship. The combination of these two items, large

---

[26] See the website for the Motley Fool, fool.com.

[27] Eric Bjorgen and Steve Leuthold, "Corporate Insiders' Big Block Trades," *Journal of Technical Analysis* (Winter-Spring 2002). This study was written in May 1998 and won the 1999 Charles Dow Award.

transaction size and privileged information, seems analogous to the positions of commercial traders in the *Commitment of Traders* report published weekly by the CFTC. Earlier in this chapter it was shown that commercials tend to be on the right side of the market while the small speculator is not.

FIBONACCI AND HUMAN BEHAVIOR

Even though we may not understand the cause underlying a particular phenomenon, we can, by observation, predict the phenomenon's recurrence.

—R. N. Elliott[28]

History is a record of great achievement in the face of disbelief, as exhibited by the explorations of Columbus, Magellan, and Marco Polo; the science of Leonardo da Vinci, Galileo, and Copernicus; and the philosophy of Socrates, Aristotle, and Plato. We are more observant today and less apt to condemn those who delve into areas still unknown. Of these, astrology is the most popular, with a very large following, particularly in Asia. Its acceptance may be partly because of its strong basis in physical phenomena. It attempts to classify personality traits based on positions of planets and stars at the time of birth, and predict the actions of groups based on the relationships of planets, moons, and comets to one another. The science of physics confirms that the positions of our moon and planets, the energy given off, and the gravitational phenomenon are directly responsible for physical occurrences of tides and weather—it seems reasonable that they might have a measurable effect on behavior. This will be considered in the following sections.

Let us look first at the fascinating subject of symmetry in nature. Science is familiar with the symmetric shapes of crystalline substances, snowflakes, the spherical planets, and the human body. The periodicity of the universe—sun spots, eclipses, and other cyclic phenomena—is also understood, but its bearing on human behavior is not known. Work in biorhythms is only at the point of being a curiosity; the relationship of behavior to nonbiological functions, such as planetary positions, is too abstract.

In 1904, Arthur H. Church wrote about phyllotaxis, the leaf arrangement of plants,[29] showing its relationship to a mathematical series based on the works of Leonardo Pisano, commonly called *Fibonacci*.[30] This mathematical series of numbers has been attributed

---

[28]R. N. Elliott, *Nature's Law: The Secret of the Universe* (New York: Elliott, 1946), 4.

[29]A. H. Church, *On the Relation of Phyllotaxis to Mechanical Laws* (London: Williams and Newgate, 1904).

[30]In the appendices to Jay Hambridge, *Dynamic Symmetry: The Greek Vase* (New Haven, CT: Yale University Press, 1931), 141–161, there is a full discussion of the evolution of this number series within science and mathematics, together with further references.

the quality of representing human behavior. Examples have been given which appear to be interesting coincidences to some, and significant to others. However, the resulting *Fibonacci ratios* have become an important part of price analysis.

It is not certain how Fibonacci conceived his summation series. His greatest work, *Liber Abaci*, written in the early part of the thirteenth century, was not published until 1857.[31] It contained a description of a situation involving the reproduction of rabbits in which the following two conditions hold: Every month each pair produces a new pair, which, from the second month on become productive; deaths do not occur. This becomes the famous *Fibonacci summation series*

$$1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots$$

(more accurately written as 1, 1, 2, 3, . . ., where the first two 1s represent the original pair of rabbits). Each element of the series is the sum of the two previous entries.

Those who have studied the life of Fibonacci often attribute the series to his observations of the Great Pyramid of Gizeh. This pyramid, dating from a preliterary, prehieroglyphic era, contains many features said to have been observed by Fibonacci. In the geometry of a pyramid there are 5 surfaces and 8 edges, for a total of 13 surfaces and edges; there are 3 edges visible from any one side. More specifically, the Great Pyramid of Gizeh is 5,813 inches high (5–8–13), and the inch is the standard Egyptian unit of measure; the ratio of the elevation to the base is 0.618.[32] The coincidence of this ratio is that it is the same as the ratio that is approached by any two consecutive Fibonacci numbers; for example,

$$\frac{2}{3} = 0.667, \frac{3}{5} = 0.600, \frac{5}{8} = 0.625, \cdots, \frac{89}{144} = 0.618$$

It is also true that the ratio of one side to a diagonal of a regular pentagon is 0.618.

Another phenomenon of the Great Pyramid of Gizeh is that the total of the 4 edges of the base, measured in inches, is 36,524.22, which is exactly 100 times the length of the solar year. This permits interpretations of the Fibonacci summation series to be applied to time.

The Greeks showed a great fascination for the ratios of the Fibonacci series, noting that while $F_n/F_{n+1} = 0.618$, the reverse $F_{n+1}/F_n = 1.618$ was even more remarkable. They called these relationships *golden sections* and appear to have used them in the proportions of such works as the Parthenon, the sculpture of Phidias, and classic vases. Leonardo da Vinci consciously applied the ratio to his art. It has always been a curiosity that the great mathematician, Pythagoras, left behind a symbol of a triangle of Fibonacci proportions with the words "The Secret of the Universe" inscribed below.

Church, in his work on phyllotaxis, studied the sunflower, noting that one of normal size (5 to 6 inches) has a total of 89 curves, 55 in one direction and 34 in another, consecutive Fibonacci numbers. In observing sunflowers of other sizes, he found that

---

[31] Leonardo Pisano, *Il Liber Abaci di Leonardo Pisano* (Rome: Baldassare Boncompagni, 1857).

[32] Jay Hambridge, *Dynamic Symmetry: The Greek Vase*, 27–38.

Chambered nautilus   Logarithmic, or "golden" spiral
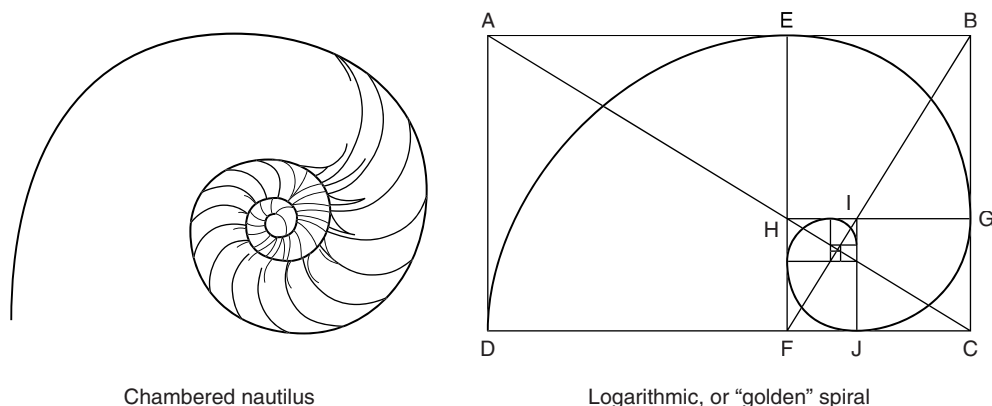
FIGURE 14.13   The Golden Spiral, also the logarithmic spiral, is a perfect representation of the chambered nautilus.

Source: Robert Fischer, *Fibonacci Applications and Strategies for Traders* (New York: John Wiley & Sons, 1993), 9. Original source: H.E. Huntley, *The Divine Proportion* (New York: Dover, 1970), iv, 101. Reprinted with permission.

the total curves are Fibonacci numbers (up to 144) with the two previous numbers in the series describing the distribution of curves. The chambered nautilus is considered a natural representation of a *golden spiral*, based on the proportions of the Fibonacci ratio (see Figure 14.13). Nature also shows that the genealogical pattern of a beehive, and the stem (growth) structure of the Sneezewort (*Achilleaptarmica*) are perfect duplicates of the Fibonacci series.

Up to now, aspects of the Fibonacci series have been intriguing, but here it goes a step beyond. The numbers in the series have frequent or coincidental occurrences:

- The human body has 5 major projections; both arms and legs have 3 sections; there are 5 fingers and toes, each with 3 sections (except the thumb and great toe). There are also 5 senses.
- In music an octave means 8, represented on the piano by 8 white keys and 5 black, totaling 13.
- There are 3 primary colors.
- The United States had 13 original states and 13 is an unlucky number.
- The legal age is 21 and the highest salute in the army is a 21-gun salute.
- The human emotional cycle has been determined at 33–36 days by Dr. R. B. Hersey.[33]
- The wholesale price index of all commodities is shown to have peaks of 50 to 55 years according to the Kondratieff wave: in 1815 after the war of 1812, 1865 after the Civil War, 1920 after the World War I, and about 1975. . . .[34]

---

[33] Elliott, *Nature's Law*, 55. Elliott quotes other human emotional relationships.

[34] *Cycles* (January 1976, p. 21); see also James B. Shuman and David Rosenau, *The Kondratieff Wave: The Future of America until 1984 and Beyond* (New York: Dell, 1974), which is based on the theory developed by the Russian economist early in this century.

These examples are not meant to prove anything in the strict sense, but to open an area that may not have previously been considered. Human behavior is not a pure science, and probes of this sort may lead the way to further understanding. The following sections deal with ideas such as these—sometimes reasonable and other times seeming to stretch the imagination.

Application of Fibonacci Ratios

Fibonacci ratios have become an integral part of charting, used to measure advances and retracements. The most popular ratio is 0.618; therefore a price advance that measures $5 from start to peak will be expected to retrace by $3.09, before starting higher again, if it conforms to this pattern. Alternately, traders have used 0.382, the complement of 0.618, as a key retracement level; however, this is not a Fibonacci ratio. Advances of 0.618 or 1.618, based on the length of the previous upwards move, are measured from the end of a retracement to predict the extent of the next leg up. Therefore, a move that retraced 0.618 can be expected to test the previous highs, or move higher by an additional 0.618 of the first leg up. Examples of these ratios can be found throughout charting analysis and they are an essential part of Elliott's wave analysis, covered in the next section. Specific examples can also be found in Chapter 4, where *retracements* are discussed.

Fibonacci ratios have also been applied to the duration of price moves by projecting the distance between peaks and valleys. This is covered in the section "Time-Goal Days."

ELLIOTT'S WAVE PRINCIPLE

R. N. Elliott was responsible for one of the more highly regarded and complex forms of market technical analysis. The *Elliott Wave Theory* is a sophisticated method of price-motion analysis and has received careful study by A. H. Bolton (1960), and later by Charles Collins. His works are fully covered in two books by Robert Prechter, who also publishes the *Elliott Wave Theorist*; brief summaries of the method appear in some of the comprehensive books on market analysis.[35] This presentation of Elliott's technique will include both the original principles and extensions, with examples.

The Wave Theory is an analysis of behavioral patterns based on mathematics and implemented using price charts; its original application was for the stock index. It has been credited with such remarkably predictive ability with respect to the Dow Jones Industrial Averages that it is second only to the recurrence of Halley's Comet. It is understood that Elliott never intended to apply his principle to individual stocks, because price movement based on relatively low activity might not be consistent with those patterns that are

---

[35]Robert R. Prechter, Jr., *The Major Works of R. N. Elliott* (Chappaqua, NY: New Classics Library, circa. 1980); and A. J. Frost and Robert R. Prechter, Jr., *Elliott Wave Principle* (Chappaqua, NY: New Classics Library, 1978).

attributed to mass behavior. This is the same justification used by Dow when he created the original Industrial Average. Caution should be exercised when applying this method to individual stocks and futures markets that might not reflect macroeconomic patterns.

The successes of the Elliott Wave Theory are fascinating and serve to reinforce the use of the technique; most summaries of Elliott's work recount them, and the reader is encouraged to read these. The *waves* referred to in the theory are price peaks and valleys, not the formal oscillations of sound waves or harmonics described in the science of physics. The waves of price motion are overreactions to both supply and demand factors within major bull moves developed in 5 waves and corrected in 3. His broad concept was related to *tidal wave bull markets* that have such large upward thrusts that each wave could be divided into 5 subwaves satisfying the same principle. After each primary wave of the major bull trend there was a major corrective move of 3 waves, which could be further divided into 3 subwaves (see Figure 14.14).

The types of waves could be classified into the broad categories of *triangles* and *ABCs*, representing a main trend and a correction, respectively. The term *triangle* was taken from
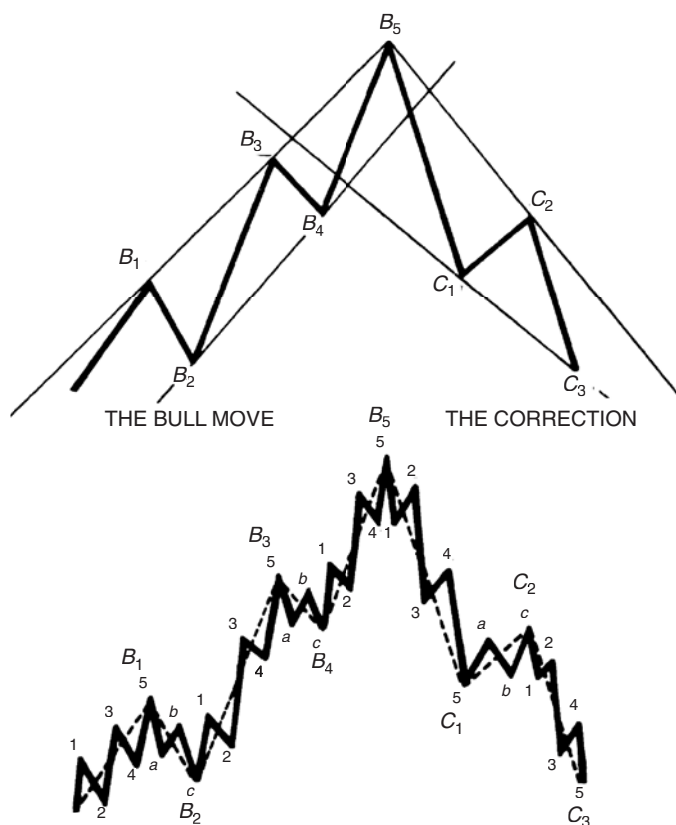
Subwaves show 5 waves in the main move and 3-wave correction

FIGURE 14.14 Basic Elliott wave.
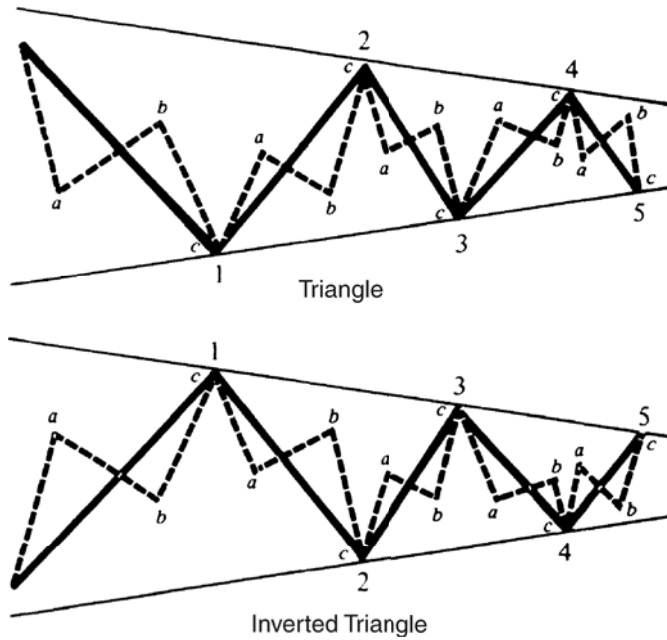
**FIGURE 14.15**   Triangles and ABCs.
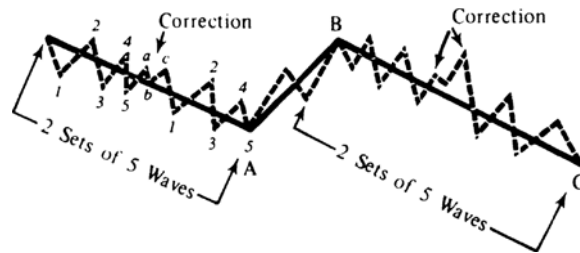
FIGURE 14.16   Compound correction waves.

the consolidating or broadening shape that the waves form within trendlines, although in later works Elliott eliminated the expanding form of the triangle (see Figure 14.15).

An interesting aspect of the theory is its compound-complex nature, by which each sequence of triangles can occur in subwaves within waves (Figure 14.16). More recent work suggests that in futures markets, a 3-wave development is more common than 5 waves. Prechter has shown many examples of major equity index moves that conform to the ratio of 1.618. The equity index, because of its great participation, is most likely to represent the idealized patterns of human behavior.[36]

---

[36]Robert R. Prechter, Jr., David Weiss, and David Allman, "Forecasting Prices with the Elliott Wave Principle," in Todd Lofton, ed., *Trading Tactics: A Livestock Futures Anthology* (Chicago: Chicago Mercantile Exchange, 1986).
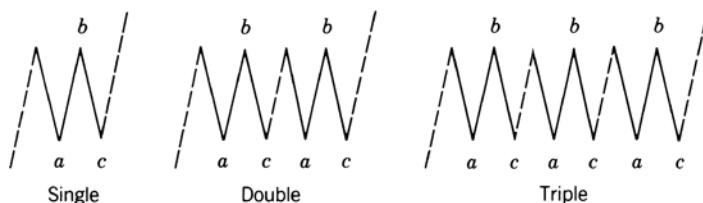
FIGURE 14.17   Elliott's threes.

Elliott's Sideways Markets

Occasionally, the market pauses during a major move; or, it may move sideways in a volatile pattern after completing the 5th leg of a wave. This has been described as "stock prices seen to be waiting for economic fundamentals to catch up with the market expectations."[37] These periods can be represented by a single *three*, a simple zigzag or flat formation, or by the more extended *double* or *triple three* (Figure 14.17).

A small variation of the *single three* has been noted to occur following the third wave, when the zigzag forms a minor swing reversal with *B* lower than its preceding top, and *C* lower than *A*. Elliott also recognized this as a *descending zigzag* in an upward trend.[38]

Fitting the Market to the Patterns

One point to remember when applying an intricate set of rules is that an exact fit will not occur often. The best trading opportunities that will arise will be for those price patterns that fit best as the move is progressing; each successful step will serve as positive reinforcement for continuing. The critical period in the identification process is the fifth wave. The failure of the fifth wave to form indicates that the last correction of three waves will be retraced. In a bull market, an extension of the fifth wave is often followed by a corrective 3-wave pattern. In addition, the recognition of a 5-wave sequence should be followed by further analysis to determine whether that cycle was part of a more complex series. One of the difficulties in this method is the orientation of the current position to the wave formation; the multitude of primary and secondary waves makes some of the situations subjective until further price movement clarifies the pattern. Anyone interested in the further interpretation of wave formation should begin with Bolton's work.

Elliott's Use of the Fibonacci Series

The application of the Elliott Wave Theory was unique in its use of the Fibonacci series. Besides the natural phenomena mentioned earlier, the summation series has the mathematical properties that

---

[37] See Prechter et al., *Forecasting Prices*.

[38] Robert R. Prechter, Jr., "Computerizing Elliott," *Technical Analysis of Stocks & Commodities* (July 1983), gives some general observations on how he would go about adapting Elliott's interpretations to a computer program.

- The ratio of any number to its successor (F_i/F_{i+1}) approaches 0.618.
- The ratio of any number to its previous element (F_{i+1}/F_i) approaches 1.618.
- The ratio of F_{i+2}/F, is 2.618.
- The two ratios $(F_i/F_{i+1}) \times (F_{i+1}/F_i) = 0.618 \times 1.618 = 1.0$

Elliott was also able to link certain measurements of the Great Pyramid to the Fibonacci series and connect the number of days in the year as well as the geometric figure of a circle to his theory. Both time and the circle will play a role in Elliott wave analysis.

While Elliott used the lower end of the Fibonacci series to describe the patterns in the stock index, it should be noted that there are increasingly larger gaps between successive entries as the series increases. To be consistent with the original principle, each gap could be subdivided into another Fibonacci series in the same manner that the waves take on a complex formation. Harrah's offers an alternate approach to filling these spaces by use of *Lucas numbers*, formed in the same way as the Fibonacci summation beginning with (1, 3) and resulting in the series (1, 3, 4, 7, 11, 18, 29, 47, 76, 123, 199, . . .). The two sets are combined, eliminating common numbers, to form **(1, 2, 3**, 4, 5, 7, **8, 13**, 18, **21**, 29, **34**, 47, 55, 76, **89**, 123, **144**, 199, **233**, . . .). The Fibonacci numbers have been set in bold type since they will receive the most emphasis, while the Lucas numbers will serve as intermediate levels of less significance. The numbers themselves are applied to predict the length in days of a price move. A bull move that lasts for more than 34 days should meet major resistance or reverse on the 55th day or on the 89th day (considering Fibonacci numbers only). It is suggested that a penetration of the 89th day should permit the series to start again with the beginning of the series added to 89 (e.g.,. . . , **94**, 96, **97**, **102**,107, **110**,118, **123**,136, . . .), including the more important Lucas and Fibonacci numbers from the original series. This effect is similar to the complex wave-within-a-wave motion.

The same numbers are used to express key levels in a trend reversal. For example, a bull move that carries prices up for 47 days before a reversal should meet resistance at the price level on the 34th day. If that price does not stop the reversal, either the behavioral implications of the number series do not hold for this situation, or prices are in a different part of the cycle.

With the introduction of Lucas numbers (L) there are some additional key ratios. In the combined Fibonacci-Lucas series (FL), denote an element with j if it is the first element of the other series, following entry i; L_A is the first Lucas number entry following Fi that is a Fibonacci number. This results in the ratios $F_i/L_j = 0.72$, $L_i/F_j = 0.854$, and $F_i/F_{i+2} = L_i/L_{i+2} = 0.382$.

The important ratio of a Fibonacci number to its following entry can be represented by the ratio of successive numbers ($\frac{1}{2}$, $\frac{2}{3}$, $\frac{3}{5}$, $\frac{5}{8}$, $\frac{8}{13}$, $\frac{13}{21}$, $\frac{21}{144}$, . . .). When expressed in decimal, these ratios approach the number 0.618 in a convergent oscillating series (1.000, 0.500, 0.667,0 .625, 0.615, 0.619, . . .). These ratios, the key Fibonacci-Lucas ratios and the alternate-entry ratios, represent the potential resistance levels, expressed as a percentage, for price adjustments within a well-defined move. For example, a price advance of 100 points in the Dow to 9,500 might correct 100%, 50%, or 62%, to 9,400, 9,450, or 9,438, respectively, according to the most important ratios.

Trading Elliott

Elliott also knew that there was great variability in the waves and ratios created by price movement. The appearance of the waves is not regular in either length or duration and should not be expected to continually increase as they develop, although the 5th wave is generally the longest. The waves must be identified by peaks only. Elliott introduced a channel into his theory in order to determine the direction of the wave being analyzed as well as to establish intermediate price objectives. Looking back at the diagram of the basic wave (Figure 14.14), note the channel that is drawn touching the peaks and bottoms of the bull move. For every two peaks, a channel can be drawn that will serve as a trendline for price objectives (see Chapter 3, "Creating a Channel with Trendlines" and "Price Objectives for Bar Charting"). A break of the lower trend line in the bull move will serve to tell when a correction has begun.

The Elliott Wave Theory is very intricate, and a full understanding requires a careful study of the original material, but the following rules outline the essentials of the method:

1. Identify a main trend.
2. Determine the current status of the main trend by locating the major peaks and bottoms that will form the 5 key waves.
3. Look for 3-wave corrections and 5-wave subtends or extensions.
4. Draw trendlines to determine the direction of the trend and reactions.
5. Measure the length of the waves in days to determine their adherence to the Fibonacci sequence, or alternatively, the Fibonacci-Lucas series; measure the size of reactions as compared to Fibonacci or Fibonacci-Lucas ratios.
6. Watch for reactions at points predicted by the Fibonacci sequence and corresponding to the patterns described by the 5-wave main trend and 3-wave correction.
7. Use the ratios, day counts, and trendlines as predictive devices to select price objectives.
8. Use the trendlines to determine changes of direction.

As complex as Elliott waves are, computer programs have been written that do a remarkably good job identifying them and showing them on a chart. To find these programs, you only need to type "Elliott wave" into an Internet search engine. These software packages have sophisticated graphics and technical analysis tools (see "Automating Elliott's Wave Analysis" below).

The Super Cycle

In all interpretive analysis, the big picture of long-term economic movement forms the foundation, and shorter-term cycles and patterns are found within that larger structure. Elliott's wave theory is no exception. Shorter time periods include relatively more noise that make frequent, erratic patterns difficult to separate from the significant market movement. In 1996, Prechter used *Supercycle* to describe the 5th wave of the prolonged bull move that was still intact at that time (see Figure 14.18), applying Fibonacci ratios to forecast targets and explain the past moves.[39]

---

[39]Robert R. Prechter, Jr., "Major Sea Change II," *Futures* (March 1996).
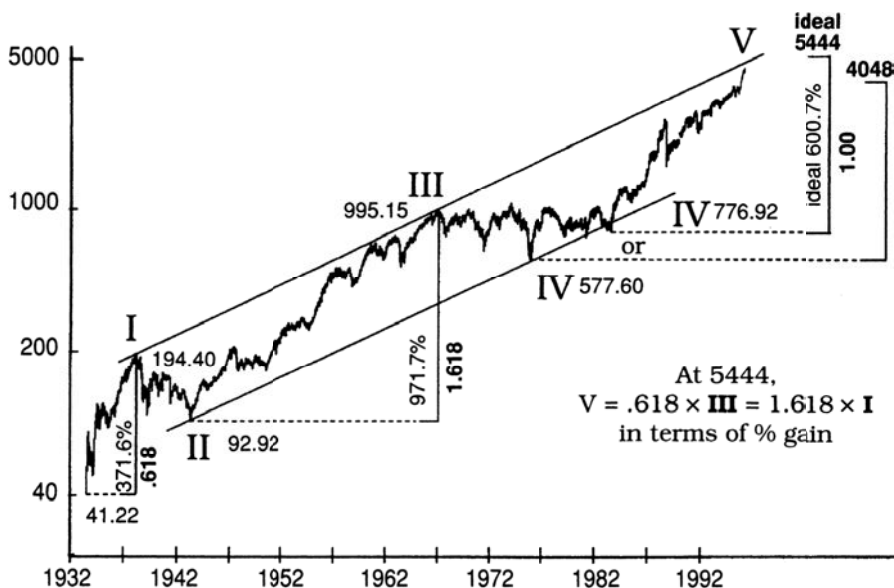
FIGURE 14.18   Price relationships in Supercycle (V).

Source: Robert Prechter, Jr., *Futures* (March 1996). Copyrighted by Robert Prechter, Jr. Reprinted with permission of *Futures*.

When using a charting technique, it is best to look for markets that, in some time frame, conform to the type of patterns you are seeking. This is also true with Elliott, which requires that prices advance in proportion to the Fibonacci ratios. In the DJIA Supercycle, shown in Figure 14.18,

- Wave 1 begins at the 1932 low of 41.22, peaking in 1937 at the high of 194.40, a gain of 371.61%, close to the Fibonacci ratio of 38.2%.
- A Wave 2 retracement follows, which is equal to 66% of Wave 1.
- Wave 3 starts in 1942 at the low of 92.29 and ends in 1966 at the high of 995.15, a gain of 970.975%. Note that when you divide the first gain by the second gain, 371.616/970.975, you get the classic Fibonacci ratio .3827. If Wave 1 is taken as equal to 0.618, then Wave 3 is 1.618.
- The Wave 4 retracement is another small move, which may be interpreted as ending in 1976 at 577.60 or, preferably in 1983 at 776.92.
- The *ideal* 5th Wave should then follow the same pattern, and gain 600.692%. This target is found by finding the value of Wave 5 such that

$$Wave\ 1/Wave\ 5 = Wave\ 5/Wave\ 3 = 0$$

This gain of 600% is based on the low of 776.92, and gives a target price of 5444. Traders are cautioned that the high volatility and unusual extension of the current move is likely to cause very specific targets, such as this, to be violated by as much as 10%.

Automating Elliott's Wave Analysis

While analysts believe that there is substantial value in Elliott's method and accept its concepts, it is an added benefit to be able to assess its success in the past. During the past few years, new programs such as *Neoware, MTPredictor*, and *Track 'n Trade* actively compete with *Advanced GET*. These programs identify complex Elliott Wave formations; however, there are simple formations that can be tested automatically with a minimum of programming effort.[40] Murray Ruggiero has defined the four trading opportunities for a basic Elliott wave pattern to be:

1. Enter Wave 3 in the direction of the trend.
2. Stay out of the market during Wave 4 (a retracement).
3. Enter Wave 5 in the direction of the trend.
4. Take the countertrend trade at the top of Wave 5.

When a wave appears in two time frames, such as in both daily and weekly charts, the likelihood of the success of this formation increases. This concept is well accepted throughout technical analysis and can be found in many sections of this book, including Chapter 19. Without confirmation of a pattern or trade, the risk increases.

The key to automating Elliott's technique is the creation of the Elliott wave oscillator (EWO). A series of extensions and pullbacks can be found using the EWO and combined with a method of counting waves that locates the current price within the standard Elliott wave pattern. The following five steps present the bull move only, although the bear market is treated as just the reverse set of rules. The full program, *TSM Elliott Wave*, can be found on the Companion Website.

1. Calculate the Elliott wave oscillator (EWO) as the difference between a 5-period and a 35-period simple moving average, applied to the average of the high and low prices,

```
            mean = (high + low) / 2
    EWO = average(mean,5) − average(mean,35)
```

2. A new upwards trend begins when the EWO makes a new high for period n, where n is determined by the user. This allows, for example, Wave 3 to be identified when it goes above the high of Wave 1.

3. A new upwards trend also begins when the current value of EWO is below zero (the 5-day average is below the 35-day average), and the trend is down (Step 2), but the EWO has rallied by a predetermined percentage (called the *trigger*) of the lowest oscillator value of the past n periods (*lowest* (EWO,n)). Then the trend is up if EWO < 0, the previous trend value is down, and

```
        EWO > trigger*lowest(EWO,n)
```

---

[40]The content of this section is based on Murray A. Ruggiero, Jr., "Building the Wave," *Futures* (April 1996). He credits Tom Joseph for the development of the computer software, which is the basis for the product *Advanced GET*.

This rule allows the next wave to be found based on the retracement that comes between each wave.

4. To relate this to Elliott, we must know where prices are located in the 5-wave sequence. For an uptrend, this is done in the following order:

   a. When the trend turns from down to up, we assume it is the beginning of Wave 3; we save the current values of EWO and the price.

   b. Continue to save the new high EWO value and the new high price for Wave 3.

   c. Wave 4 begins when EWO falls to zero.

   d. If Wave 4 is currently active and the price is a 5-period high and EWO > 0, then Wave 5 begins. Save the highest EWO value and highest price of Wave 5 whenever they occur.

   e. If EWO in Wave 5 becomes higher than the highest EWO value in Wave 3, then we are still in Wave 3. Relabel the current Wave 5 values to be Wave 3, and continue as in Wave 3 with Step 4b.

   f. If the trend turns from up to down in Wave 5, then this is a *Wave 3 down*. Reset all values and look for a new *Wave 3 up* when Step 4a is satisfied.

5. The trading rules needed are based on the values of n, the number of periods, and the *trigger* (the percentage retracement).

   a. *Buy* on the open when Wave 3 is first identified, whether on a new n-period high or a retracement of the previous downturn (Step 4a).

   b. *Buy* on the open when Wave 5 is first identified (Step 4d).

   c. *Buy* on the open when Wave 5 turns into Wave 3 (Step 4e).

   d. *Close out* any long position when EWO falls below zero.

With these rules, a long position is taken during the strongest parts of Wave 3 and Wave 5. As a guideline for selecting both the period and the trigger values, a test of the EURUSD futures from 2000 through May 2011 showed excellent results using a trend period of 40 days, oscillator periods of 5 and 35 days, and a retracement *trigger* value of 0.35. These parameters were found to be the best for the S&P over that same period. Figure 14.19 shows the most recent trades, entering in September 2010, exiting in November, and reentering in January 2011. The programs for the functions, indicators, and strategies (beginning with *TSM Elliott Wave*) can be found on the Companion Website. Additional programs for the Elliott wave count, including trading rules, can be found in Ruggiero.[41]

A Comment on Elliott Wave Theory

The Elliott Wave Theory is highly regarded, although it is an intricate combination of mathematics and chart interpretation based on assumptions about human behavior. Because it is primarily based on chart patterns it has been criticized as being too interpretive. The development of the 5th wave, the grand finale, receives the most attention— sometimes it never

---

[41] Murray A. Ruggiero, Jr., *Cybernetic Trading Strategies* (New York: John Wiley & Sons, 1997).
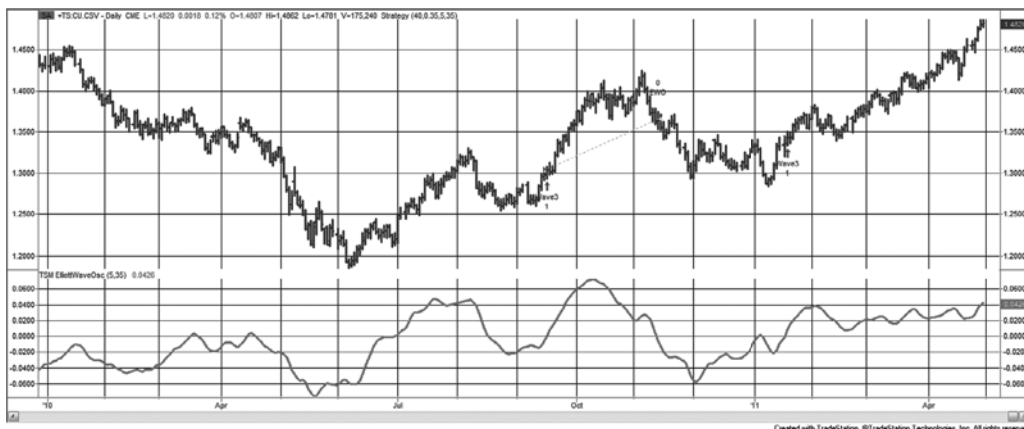
FIGURE 14.19   Most recent trading signals for the Elliott wave strategy, and the Elliott wave oscillator, applied to EURUSD futures.

develops, and at other times it must be extended into another subset of waves. An analysis by Merrill shows that the median stock market bull move has 7 legs and the bear move has 5 legs, which may account for the frequent use of the compound form of the Elliott wave. Those analysts who find this study of interest should also read the works of W. D. Gann and Edson Gould, both of whom concentrated on mathematical approaches to charting.

PRICE TARGET CONSTRUCTIONS USING THE FIBONACCI RATIO

Harahus shows interesting constructions using Fibonacci ratios; these are referred to as the *golden rectangle*, *golden triangle*, and *golden spiral*. Although there is no doubt of their importance, their application to markets can be complicated. In using a method as sophisticated as the Elliott Wave Theory, it is necessary to select situations that are most representative of the phenomenon described by the Fibonacci series. Experience will help you decide which waves and which formations are most important and should be applied to this geometric analysis.

Harahus introduced the *regular pentagon*, which has sides of equal length, as a tool for measuring corrections to the primary waves. This pentagon has the property that any diagonal is 1.618 times the length of a side, exactly a Fibonacci ratio. By constructing a regular pentagon so that the major price trend falls along one diagonal, or along one side, the other lines connecting the corners of the pentagon will serve as support of, or resistance to, future price moves. In addition, the circumscribed and inscribed circles will serve as measurements of support and/or resistance (Figure 14.20). The use of a circle is similar to Gann's concept of *time and space*, which will be covered a little later in this section.

Harahus extends the charting techniques of the Elliott Wave Theory using circles and arcs, a method that has been applied to measuring retracements. A circle drawn from the top or bottom of a wave, representing the 38, 50, or 62% levels, serves as a convenient measurement of elapsed time combined with a price correction. Prices are expected to meet resistance at any attempt to penetrate the key circles formed using the center points *A* or *B* in Figure 14.21.
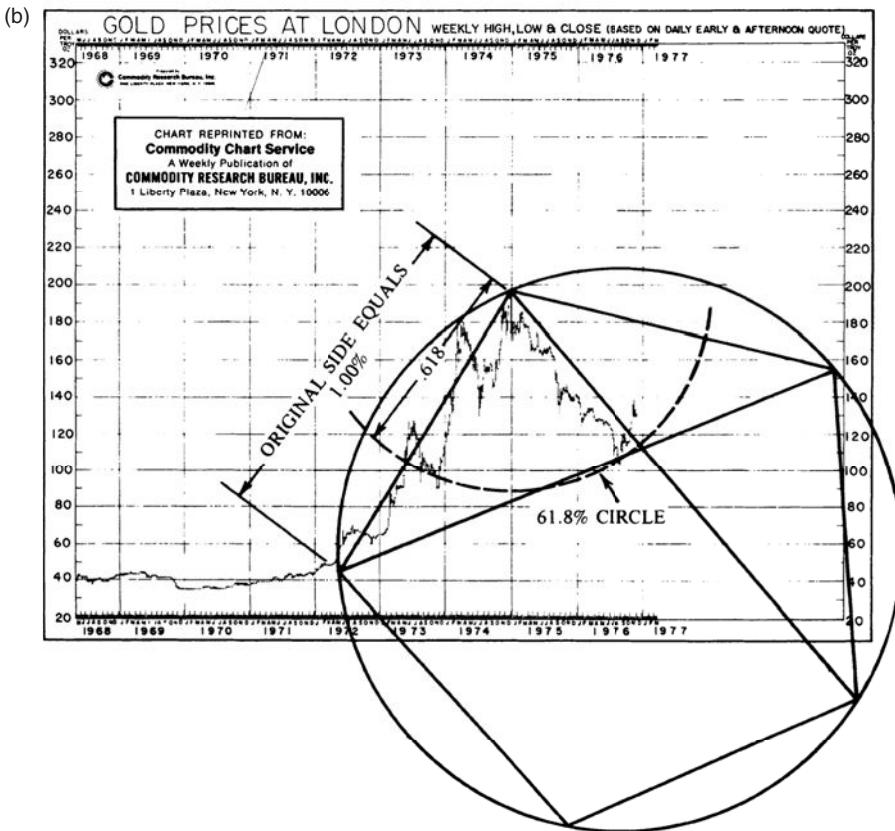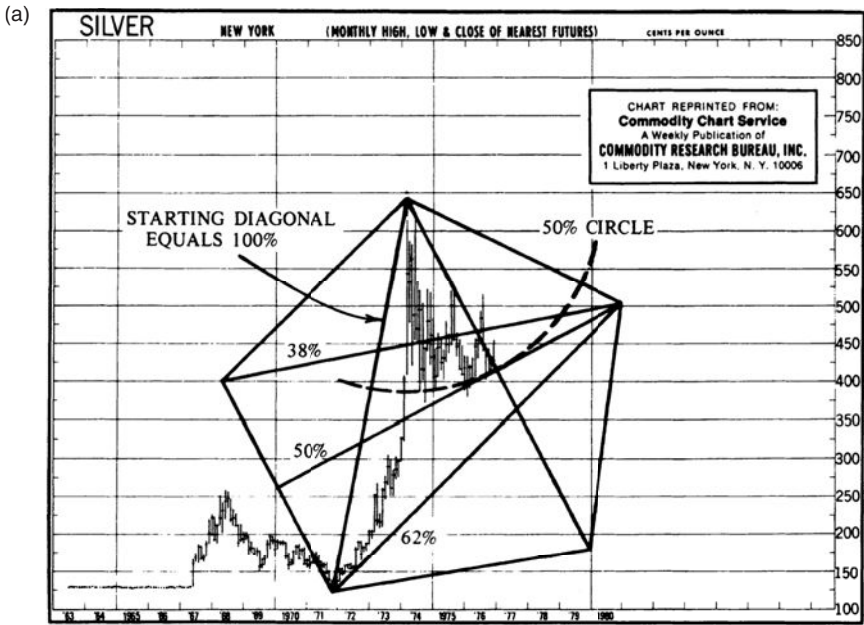
FIGURE 14.20 (a) Pentagon constructed from one diagonal. (b) Pentagon constructed from one side.
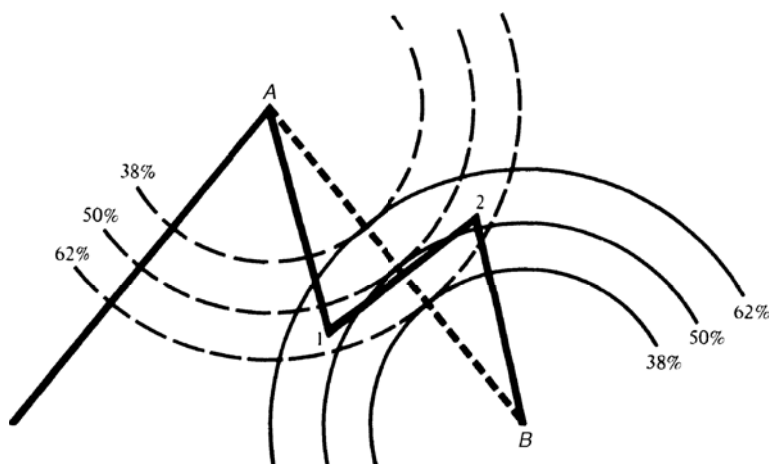
FIGURE 14.21   Using circles to find support and resistance.

Alternate Arc Measurement

In Chapter 4, there were alternate ways of measuring retracement levels. Tom DeMark[42] identified the extent of a price move differently from the classic definition. In his approach, if the market is currently at a low, rather than projecting the distance of this drop from the most recent swing high, he chose to look for the highest point that had occurred since the last time the market traded at this low level. He then applied his choice of ratios to find the key retracement levels. DeMark also uses his measurement from high to low to draw arcs identical to those shown in Figure 14.21 but limited to the Fibonacci ratio 0.618 and its complement 0.382.

FISCHER'S GOLDEN SECTION COMPASS SYSTEM

Although there are many systems that use the Fibonacci ratios, Fischer's *Golden Section Compass* (GSC) *System* is the only one based primarily on that concept and fully automated.[43] The system is founded on the premise that human behavior is reflected in the Fibonacci ratios; the human decision-making process unconsciously selects points to act that appear to be the right time or the right price level, but turn out to coincide with the same points determined by Fibonacci ratios. The behavior-based rules are combined with practical entry and stop-loss rules to eliminate those situations that do not properly develop.

---

[42]Thomas DeMark, "Retracing Your Steps," *Futures* (November 1995). Also see Chapter 2 of DeMark, *The New Science of Technical Analysis* (New York: John Wiley & Sons, 1994).

[43]Robert Fischer, *The Golden Section Compass Seminar* (Hamilton 5, Bermuda: Fibonacci Trading, Ltd.,). Also see Robert Fischer, *Fibonacci Applications and Strategies for Traders* (New York: John Wiley & Sons, 1993).
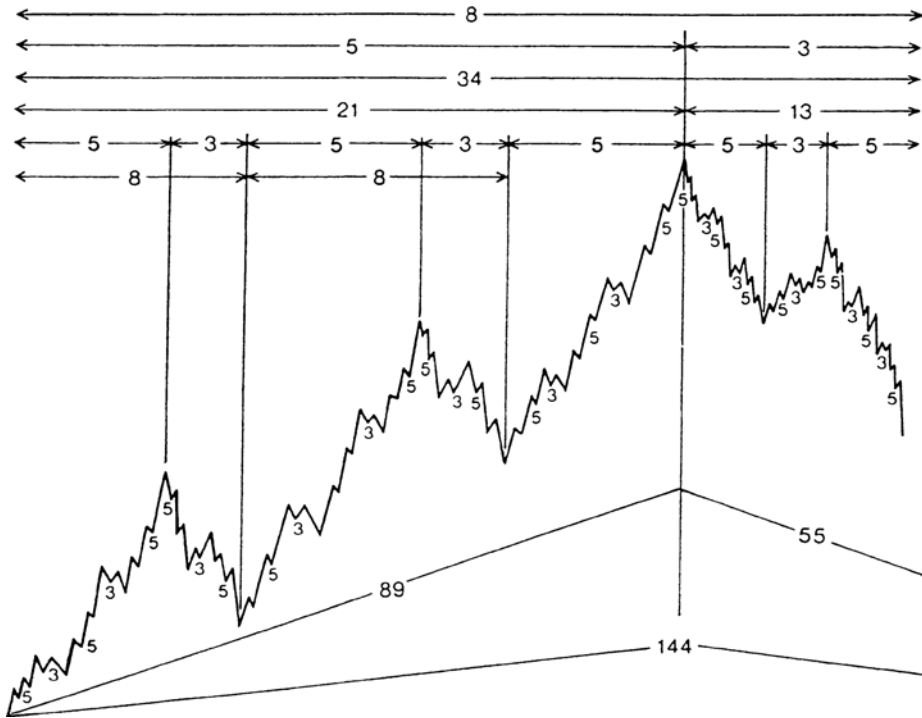
FIGURE 14.22   Elliott's complete wave cycle.
Source: Fischer, *Golden Section Compass Seminar*, 28.

Fischer is not alone in his observation of these patterns. Elliott, discussed in the previous sections, defined market cycles in 5-wave patterns with 3-wave corrections (Figure 14.22). In *Nature's Law*, Elliott shows the *complete market cycle*[44] as:

Number of	Bull Market	Bear Market	Total
Major waves	5	3	8 complete cycles
Intermediate waves	21	13	34 complete cycles
Minor waves	89	55	144 complete cycles

Time-Goal Days

The GSC System states that a new price direction will begin on, or shortly after, the day calculated as:

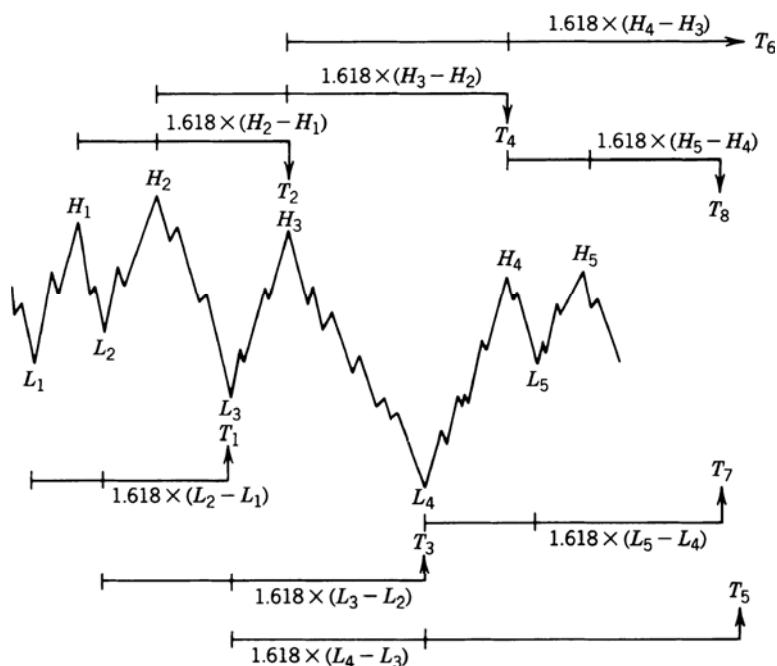$$T_k = 1.618 \times (L_i - L_{i-1}) + L_i$$

or

$$T_{k+1} = 1.618 \times (H_i - H_{i-1}) + H_i$$

where L_i and H_i are the days on which lows and highs occurred, and L_i occurs before H_i.

---

[44] Fischer, *The Golden Section Compass Seminar*.

FIGURE 14.23 Calculation of time-goal days.

For simplicity, an extreme may be used only twice for a calculation—once as the first point and once as the second. Figure 14.23 shows the order in which the calculations occur. Time-goal day T_5 was calculated from lows L_3 and L_4 before time-goal day T_7 occurred. When more than one time-goal day occurs at the same point, or when the highs or lows that formed the time-goal days are more significant, the likelihood of a major reversal increases.

Specific entry signals occur on a 5-day price reversal. A long signal is given following a time-goal day when the closing price is higher than the high of the past 5 days. This is sufficient to identify a trend change and allows the stop-loss to be the recent low or a fixed point below the entry price.
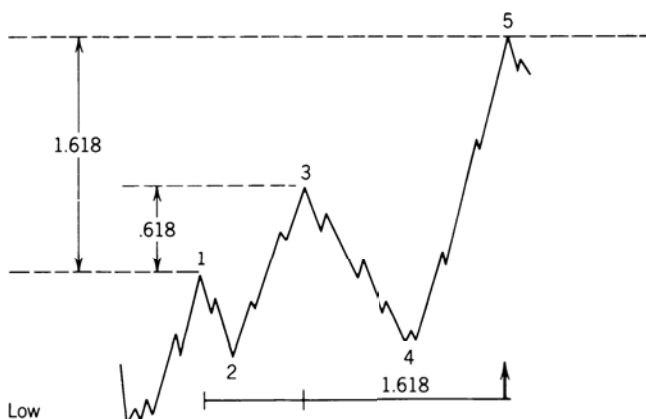
Price Goals

Price objectives are determined using the same highs and lows that were applied to time goals. Elliott's Wave Theory is refined by Fischer, as shown in Figure 14.24. Once the low has been found, followed by Wave 1, the price objectives for Waves 3 and 5 can be calculated as a function of Wave 1:

Wave 3 objective $= 0.618 \times$ (High of Wave 1 – Low of Wave 1) + High of Wave 1

Wave 5 objective $= 1.618 \times$ (High of Wave 1 – Low of Wave 1) + High of Wave 1

The pullbacks to points 2 and 4 in Figure 14.24 are not determined in this calculation, and the probability of the 5-wave target being successful is not confirmed unless the Wave 3

FIGURE 14.24   Price goals for standard 5-wave moves.

objective is satisfied. Price goals are used for profit-taking rather than new entry points. The GSC System, however, does not require that markets move in 5-wave patterns. Price objectives can be calculated from long-term lows and the high of the first move up by multiplying that move by the ratio 1.618. When price and time goals occur at the same point, there is greater confidence in the signal.[45]

In general, retracements from sustained moves can be expected to approximate or exceed 38.2% of the initial move $(1 − 0.618$, the complement of 61.8%). Once a high-level objective has been reached and a reversal occurs, consistent profits may be taken using this goal. If prices fail to make new highs following the first retracement, objectives for the next lower levels down can be set according to the inverted Elliott wave patterns, or *1.618 × initial downward move.*

Filtering Highs and Lows

The GSC System can be made to identify more significant highs and lows by increasing the *selection filter*. For example, a sensitive system would select highs and lows in gold that are separated by a minimum swing of $10 or $20 per ounce. More significant points may be identified by swings of $50 per ounce. It can be demonstrated that a small filter will generate highs and lows that are produced by noise rather than significant behavioral actions. The system cannot be forced to produce more signals than the natural patterns allow. Readers will find this the same as the concept of a *swing filter*.

The *selection filter* should be adjusted for market volatility. At higher price levels, the noise also increases and will obscure the more important high and low points. Although Fischer only briefly touches on this, traders who apply this method should be prepared to adjust the selection filter from time to time or to create a continuous volatility-adjusted filter.

---

[45]Examples of time and price objectives can be found in Tucker J. Emmett, "Fibonacci Cycles," *Technical Analysis of Stocks & Commodities* (May 1983 and March-April 1984).

W. D. GANN—TIME AND SPACE

The works of W. D. Gann cannot be explained with any thoroughness in a few words, but some of his main ideas have been selected and presented in this section.[46] Gann was a pure technician, using charts for all his analyses. His methods varied substantially from conventional charting techniques, but his philosophy was one of a professional trader: *Conserve capital and wait for the right time.* Gann traded primarily grains for many years, and in his writings he attempted to summarize his most important observations; some of them are reminiscent of other well-known market lore.

Price moves are never exact. Gann was a believer in support and resistance lines, but expected some violation of the objectives because of *lost motion*, his way of accounting for the momentum that carries prices higher or lower than their fair value. Nearly a cross between Elliott's *waves* and Angas' *cycles*, Gann classifies bull and bear moves into four *stages*, each one compared to a trending move and a subsequent reversal culminating in a major top or bottom. He observed that bull markets last longer than bear markets. He concluded that reversal patterns must decline in magnitude as the move develops and persists. A similar argument is expressed in the *theory of contrary thinking.* Much of Gann's work is based on the value of 1940 cents and requires an economic inflator close at hand to adjust prices to today's levels.

Gann's techniques combine mathematics and geometry with time and space; he finds duration as important as the size of the price change. One of his principles reflects the idea that a longer consolidation period results in a longer price move after a breakout. One popular approach to price objectives in bar charting is exactly this idea.

Time and Price

Gann proposed certain natural divisions for price swings, expressed as percentages. Zero and 100% are the most important of these. Using his understanding of human behavior, he considered a potential resistance level to be 100% of the original point of the move, or twice the starting price, or the highest or lowest price previously recorded for that market (based on cash prices, not futures). In a reversal, 100% was a full retracement of the original move back to its starting point. Part of the rationale for this theory is that most traders like whole numbers; for this reason, orders in grains are most often placed at 5- and 10-cent levels rather than at odd values. Even now, traders will find more activity as prices attempt to cross whole number barriers, such as increments of 100 for the S&P, or each 1,000 for the Dow. Most investors rarely think of entering an order at an odd price.

After the 100% level, decreased importance goes to increments of 50%, 25%, 12½%, and so on. For the grain markets, which were the most active at the time of Gann's trading, this would mean that major resistance could be expected at the even dollar levels with the next resistance at 50¢ intervals, then every 25¢, and so on; after a bull move of $1 per bushel, the major support would be $1 lower, then 50¢, 25¢, and 75¢, and so on. The use of successive

---

[46]William Gann's original works, *Wall Street Stock Selector* and *Truth of the Stock Tape*, both published in 1930, can be found in reprints by Lambert-Gann.

halving of intervals was also extended to time. Considering a year to be a full cycle of 360°, a half year is equal to 26 weeks, a quarter year 13 weeks, an eighth year 45 days, and a sixteenth year 22½ days. In cases of conflict, time always took precedence over price. For grains, a 1-year cycle was significant because seasonality is implicit in their supply and demand, and their price. The combination of a key price level, expressed as a percentage move, occurring at a periodic time interval is the basis for much of Gann's work.
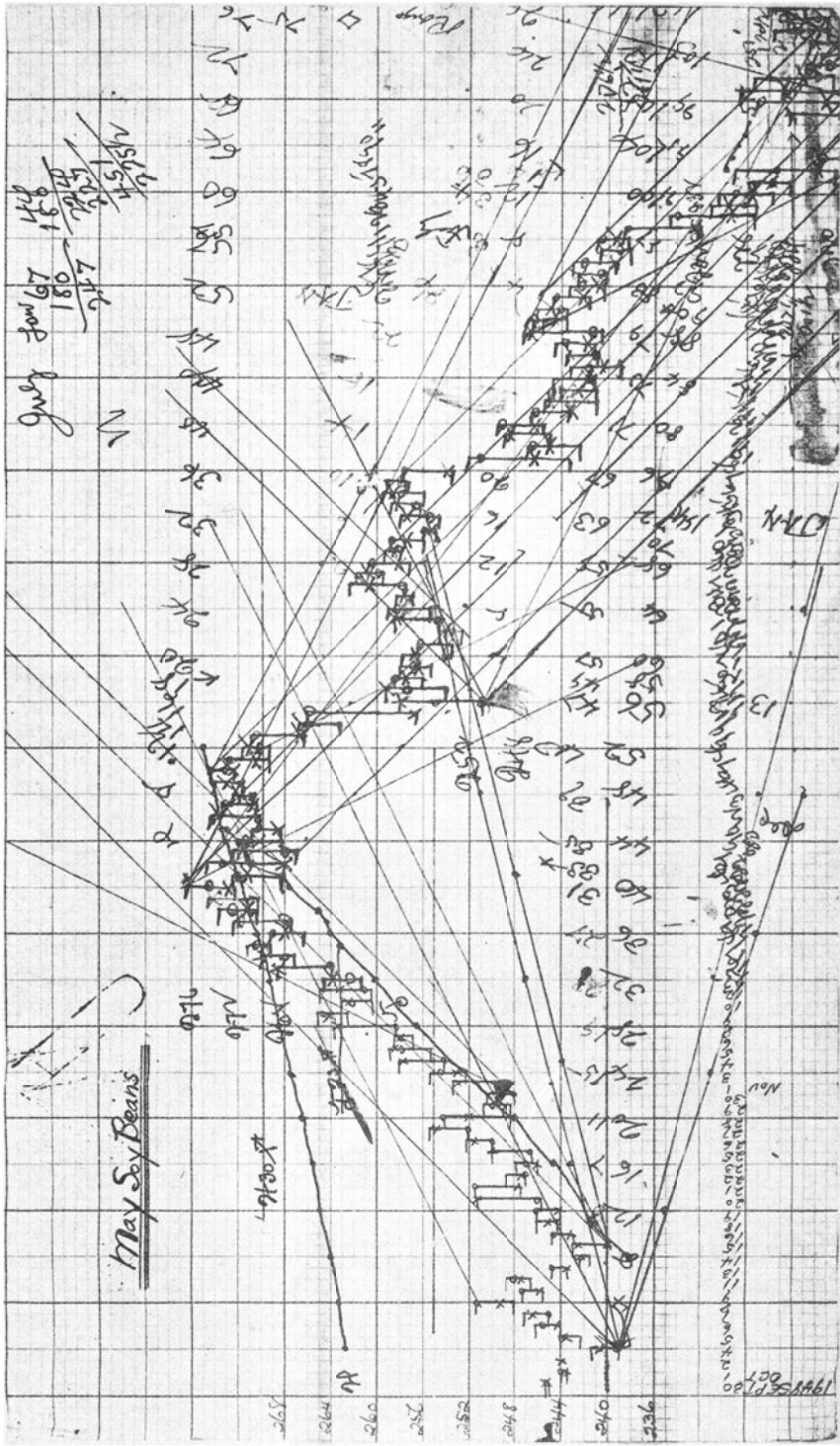
Geometric Angles

The most popular of Gann's methods is his use of geometric angles for relating price and time. By using square graph paper, it was not necessary to know the exact angle because the construction was based on the number of boxes up versus the number of boxes to the right. A 1×1 angle (45°) was drawn diagonally from the bottom of the lowest point of a price move through the intersection 1 box up and 1 box to the right. This is the *primary bullish support line.* The *primary bearish resistance line* is drawn down from left to right from the highest price using the same 1×1 angle. The next most important angles in order of significance are 2×1, 4×1, and 8×1, equivalent to about 63°, 76°, and 83°; for lower support areas, there is also 1×2, 1×4, and 1×8, or 26°, 14°, and 7°. Places where the support and resistance lines cross are of special significance, indicating a major congestion area.

Figure 14.25 is taken from Gann's private papers and shows the use of geometric angles in an actual trading situation. Lines were first drawn where Gann expected a bottom, then redrawn as prices developed. The initial upward move followed the primary 45° line; the second important support line, 1×2, met the primary downward line at the point of wide congestion at the center of the chart. The highest point on this congestion phase became the pivot point for the next 45° downward angle defining the next breakout. Traders have found that the primary 1×1, or 45° line, is an important tool for staying with the major trend. It is used to filter out small reversals in both standard charting techniques and the point-and-figure method.

Gann combined this approach with a more remarkable technique, the *squaring of price and time.* It was fortunate to find a chart that complemented Figure 14.25, based on the lowest recorded cash price of soybeans, 44¢ per bushel. Figure 14.26 shows how Gann constructed this square, beginning with the lowest price at the center and moving one square to the right, circling counterclockwise and continuing the process. The basic geometric lines (horizontal, vertical, and diagonal) indicate the major support and resistance price levels; the most important one being 44 at the center, the junction of all lines.

Relating the square to Figure 14.25, the price chart showing geometric lines, the first support level is seen to be at exactly 240 (upper left diagonal), the major resistance at 276 (right horizontal), the next minor support at 268 (lower right diagonal), congestion area support at 254 and 262 (1 box off), and back down to support at 240. Notice that the distances between the lines on the square become wider as prices increase, conforming to the notion of greater volatility at higher prices. It is also expected that soybeans at $10 will have some *lost motion* near these key support and resistance levels.
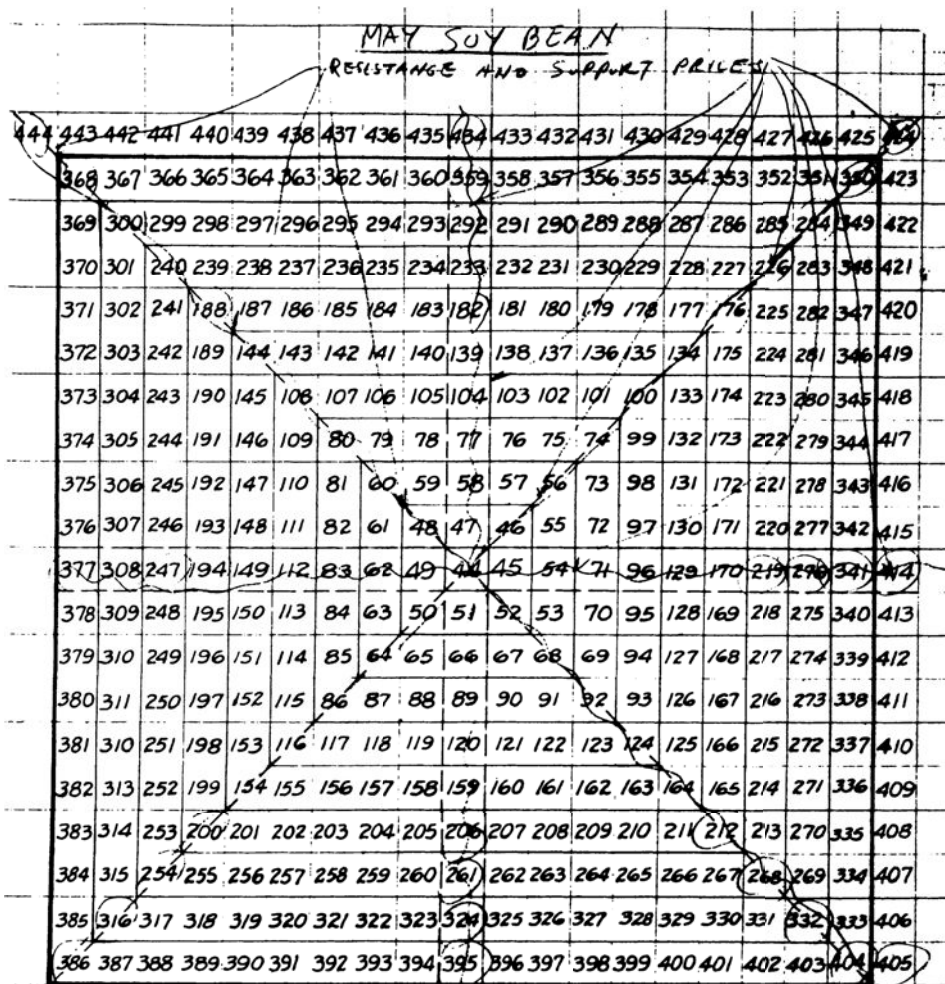
FIGURE 14.25 Gann's soybean worksheet.

FIGURE 14.26   May soybean square.

The Hexagon Chart

Gann extended his *squaring method* to include both geometric angles and the main cyclic divisions of 360°. By combining these different behavioral concepts, the strongest levels of support and resistance are found where all three coincide. The generalized construction for this purpose is the *Master Calculator*, based on aligning the chart at a point representing a multiple of the lowest historic price for that market; crisscrossing angles will then designate support and resistance for the specific market. Other time charts of importance are the *Square of Twelve* (one corner of the Master Calculator), the *Hexagon Chart*, and the *Master Chart of 360°*. The Hexagon Chart can be used as an example of the combined effect.

As shown in Figure 14.27, the inner ring begins with six divisions, giving Gann the basis for the chart name. Each circle becomes larger as it proceeds outward. In using the
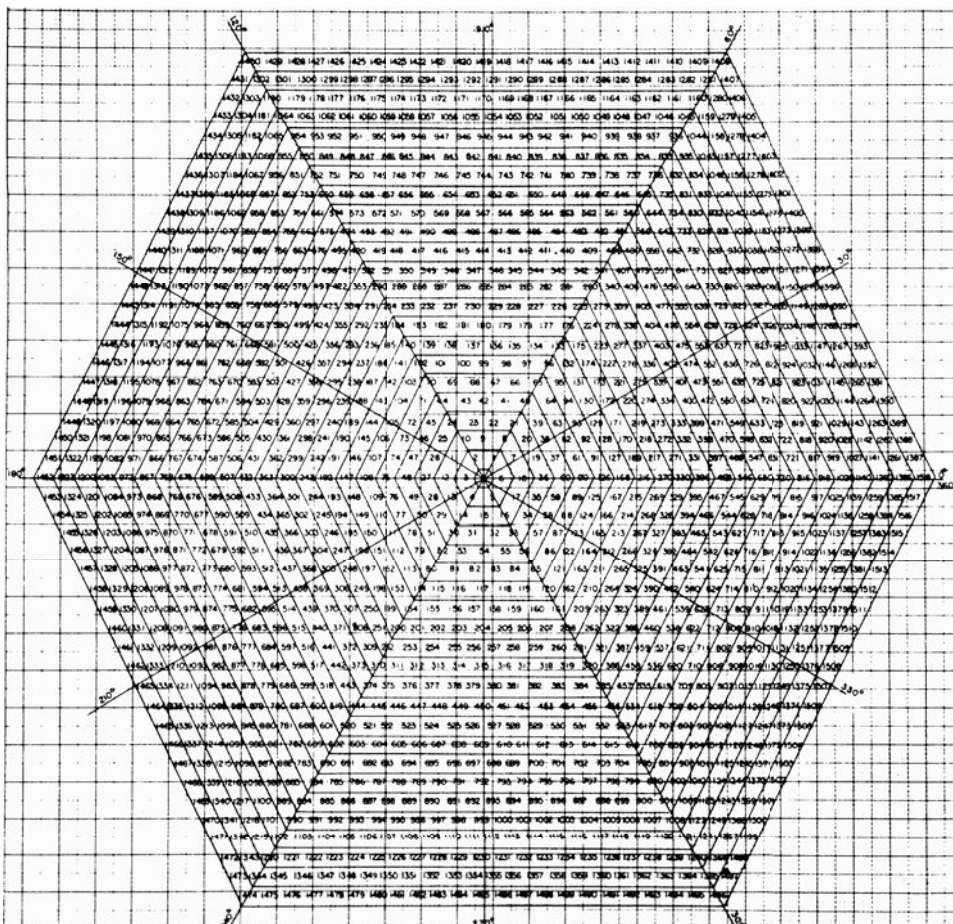
FIGURE 14.27 The Hexagon Chart.

hexagon, the degrees represent time and the numbers in the circle are price; a major support or resistance point exists when both time and price occur simultaneously.

For example, consider the 360° of the hexagon relating to the calendar year, or perhaps the crop year for grains. In his own work on grains, Gann equated 0° to March 20, near to the first day of spring, the *spring equinox*, when the sun crosses the equator going north. Then, the 45° line is on May 6, 90° on June 21 (the *summer solstice*), 180° on September 23 (the *fall equinox*), and 270° on December 21 (the *winter solstice*). These primary divisions also represent the most significant places for price support and resistance. The other lines represent secondary levels.

When looking at price and time together on the Hexagon Chart, the distance between the major degree lines becomes greater as prices increase, again showing the importance of volatility. Using the price of November 1977 soybeans, the chart shows that between 90° and 180°, or June 21 to September 23, 1977, the price of soybeans should have support at 567. Then it should move its major support level to 507 and its major resistance to

588, with the next higher and lower support and resistance at 432 and 675, respectively. As it turned out, this was a very accurate prediction.

Mathematically, when you divide the value in an outer ring with the value in the next ring towards the center of either the square or hexagon chart, we get a similar percentage move. If we take the top values of the square in the center, declining towards the middle of the chart, we get 434, 359, 292, and 233. The percentage changes are 20.9, 22.9, and 25.3, not the same, but interesting. These numbers are less consistent toward the starting value in the center of the chart.

On many of Gann's charts there is notation showing planetary movement, not related to the cycles of seasonality, but what is believed to be the Jupiter-Saturn cycle, discussed in the next section. These techniques make Gann's work more difficult to reproduce than most methods; his tools are less conventional than others. If Gann were asked for a word of advice, there is no doubt that he would caution to patience, stating: *When price meets time, a change is imminent.*[47]

Gann Software

As with the Elliott Wave, sophisticated software is now available to chart Gann patterns. Some of the software includes both Elliott and Gann, in addition to elements of financial astrology discussed in the next section. Current software includes *GannTrader 3*, *Market Analyst*, *Gannalyst*, *WaveTrader*, and *Gann Forecaster*; however, scanning the Internet for "Gann trading software" will give you an up-to-date list.

FINANCIAL ASTROLOGY

Astrology seeks a common bond in human behavior, similar to the work in biological rhythms and cycles. The impact of astrology on civilization has been great; observations of the periodicity of the moon are traced back 32,000 years. Star charts were known to have been used in Egypt about 4200 B.C., and the earliest written ephemerides were in the seventh century B.C.[48] The pyramids at Gizeh are said to have sloping corridors leading from the faces to the interior that were used as sighting tubes for Egyptian astrologers for making accurate forecasts. The acceptance of astrology throughout history is widespread, including virtually all civilizations.

Even sophisticated analysts have been found to confuse *astrology* with the daily horoscope found in a local newspaper. *Astrology*, the interpretation of the effects of planets and stars on human affairs, is an art followed by a large portion of the world population and should not simply be discarded as mystical or occult. Interpretation of these effects can be complex and involves special skills. Most forecasts begin with a birth chart, which

---

[47] Computer software is available for calculating and plotting much of Gann's works. See "Ganntrader I," *Technical Analysis of Stocks & Commodities* (January/February 1984), or contact Gannsoft Publishing Co., 311 Benton St., Leavenworth, WA 98826, or Lambert-Gann, P.O. Box O, Pomeroy, WA 99347.

[48] Derek Parker and Julia Paricor, *The Compleat Astrologer* (New York: McGraw-Hill, 1971), 12.

describes the position of the stars at the time of the inception of an event and then look at the current positions to identify the transitions. Over time, planets have taken on an association with specific raw goods, and constellations are associated with certain types of business; these relationships can now be verified using computer programs that check the intricate position of the planets with the financial statistics of companies to find correlations. However, that task will be left for the more ambitious.

We must all agree that the positions of the planets and sun—those bodies with the largest gravitational pull on the Earth and each other—have clear physical effects that we can see around us. In the study of physics, this relationship is given as

$$Force\ of\ attraction = \frac{Gravitational\ constant \times Mass}{Distance\ between\ the\ centers\ of\ the\ two\ masses}$$

where the gravitational constant is 6.67×10^{-8}, the mass is found in Table 14.3, and the distance between the centers is calculated from the Distances in Table 14.3.

Of the most obvious physical effects, the seasons and the tides are undeniable; and there are many more subtle phenomena when you study astronomy, the science of the motion of the stars and planets. In this section, we will concentrate our attention on the physical phenomena of planetary motion that can be identified and tested with the same confidence, and perhaps more, as other trading systems. Among these phenomena, eclipses and the lunar cycles are the most dominant, but there is one other combination that is important, the Jupiter-Saturn cycle.

The Jupiter-Saturn Cycle

Jupiter and Saturn combine to represent the overwhelmingly largest mass in our solar system other than the Sun (see Table 14.3). The two planets amount to about 95% of the total mass of all the planets, and are about 0.7% of the mass of the Sun. Because they are in adjacent orbits, they have a very large gravitational pull on other planets when they are near one another and a very different effect when they move apart. They are so large that they cause the Sun to shift periodically based on their positions around the true center of mass of the solar system, called the *barycenter*. When Jupiter and Saturn are on opposite sides of the Sun, the center of mass is near the center of the Sun, but when the two large planets are together, the Sun is pulled away from this center. This positioning also has significant effects on the Earth's climate, which in turn affects agricultural production, supply and demand, and ultimately the economy.[49] The Greeks called the Jupiter-Saturn cycle *The Great Maker of Time*, and it is said that every significant cycle in stocks, commodities, and interest rates is either a multiple or a harmonic of the Jupiter-Saturn cycle. The most well-known, the 59-year *Kondratieff cycle*, is 3 times the Jupiter-Saturn cycle ($3 \times 19.859 = 59.577$). Gann, who used a combination of time and price, based time

---

[49]For a complete discussion of this topic, see Henry Weingarten, *Investing by the Stars* (New York: McGraw-Hill, 1996). A section written by Richard Mogey, "Long Cycles and the Master Time Factor," is the basis for the next section.

TABLE 14.3 Size and Position of the Planets and Earth's Moon

	Equatorial Diameter (km)	Mass* (kg)	Distance from Sun (mil km)	Duration of Orbit (Earth years)
Sun	1,392,000	1.99×10^{30}	0	0
Mercury	4,880	3.34×10^{23}	57.9	0.24
Venus	12,102	4.87×10^{24}	108.0	0.62
Earth	12,756	5.98×10^{24}	149.6	1.00
Mars	6,794	7.35×10^{22}	227.9	1.86
Jupiter	142,800	8.39×10^{27}	778.3	11.86
Saturn	120,000	5.00×10^{27}	1,427	29.5
Uranus	52,000	3.77×10^{26}	2,870	84
Neptune	47,500	3.29×10^{26}	4,497	165
Pluto	2,500	5.98×10^{22}	5,893	248
Earth's moon	3,476	7.34×10^{22}	384,400 km	0.075

*Note that *mass = volume × density* and, except for the Earth and the Earth's moon, the density of the planets can only be estimated.

expectations on seasonality and cycles. His work referred to a *Master Time Factor* that was never defined, but experts believe that the Jupiter-Saturn cycle fits his technique.
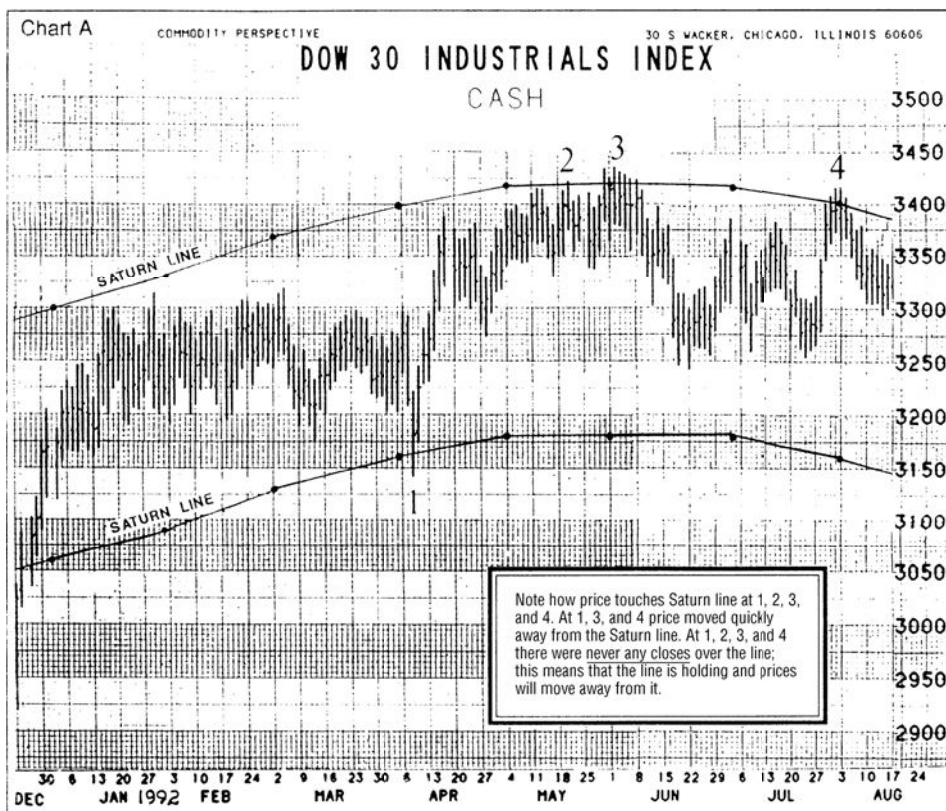
Charting the Saturn Line

Having decided that the movement of the largest planets affects the way markets move, you can create a *planetary envelope* based on Saturn's path with respect to the Dow Jones Industrial Average and find frequent concurrence. This requires converting the position of the planets into price.[50] When completed, this envelope is treated as support and resistance lines in a manner similar to the standard charting interpretation, where penetration of a resistance line will become support; however, in appearance the Saturn lines are curved to represent its cycle (see Figure 14.28).

Converting Planetary Position to Price

Long transforms planetary position to price using two separate *wheels*, combined to form a *Universal Clock*, in a manner similar to W. D. Gann. Each wheel is unique to a specific market. To create the Universal Clock for the DJIA, you begin above the right horizontal (at 3 o'clock) with the number 1 and move counterclockwise, placing 6 numbers per quarter, 24 for an entire cycle. You then move to the next outer circle and continue with the number 25 just outside the original number 1; therefore, each concentric circle contains the next 24 values, ending at 360° (a full cycle). The DJIA Clock is shown in Figure 14.29.

In the outer wheel of the clock, the DJIA is shown in increments of 10 points beginning in 1930 and continuing in circles of 24 values. Long does not explain the choice of the beginning price, but Gann uses a significant low and astrologers tend to use the

---

[50]Jeanne Long, "Planetary Support and Resistance on the DJIA," *A Trader's Astrological Almanac 1994* (PAS Astro-Soft, Inc., 450-106 State Road 13 North, #206, Jacksonville, FL 32259-3863, email: pas24@aug.com).

FIGURE 14.28 Saturn lines drawn on the Dow Industrial Average.

Source: Jeanne Long, *A Trader's Astrological Almanac* (1994). Reprinted with permission from PAS, Inc., e-mail: pas24@aug.com, website: www.galacticinvestor.com.

price at a key starting time, such as the beginning of a new Dow calculation, birth of an exchange, or a transforming event.

To locate the planet on the time wheel, begin with the planet's position on a specific date. For example, on October 1, 1993, Saturn was at 24° 14′ (with respect to Aquarius). Aquarius falls between 300° and 330°; therefore, the position of Saturn is 300° + 24° + 14′, or approximately 324°. By locating 324 on the inner part of the wheel, you can refer to the prices on the outer part of the wheel and find the two prices (each in adjacent circles) that span the current DJIA price. These values are the current support and resistance lines. New values are found monthly.

Major Physical Events

Of the major physical phenomena used in financial astrology, the most important are:

- *Solar eclipse*, when the moon passes between the Earth and the Sun.
- *Lunar eclipse*, when the Earth passes between the moon and the Sun.
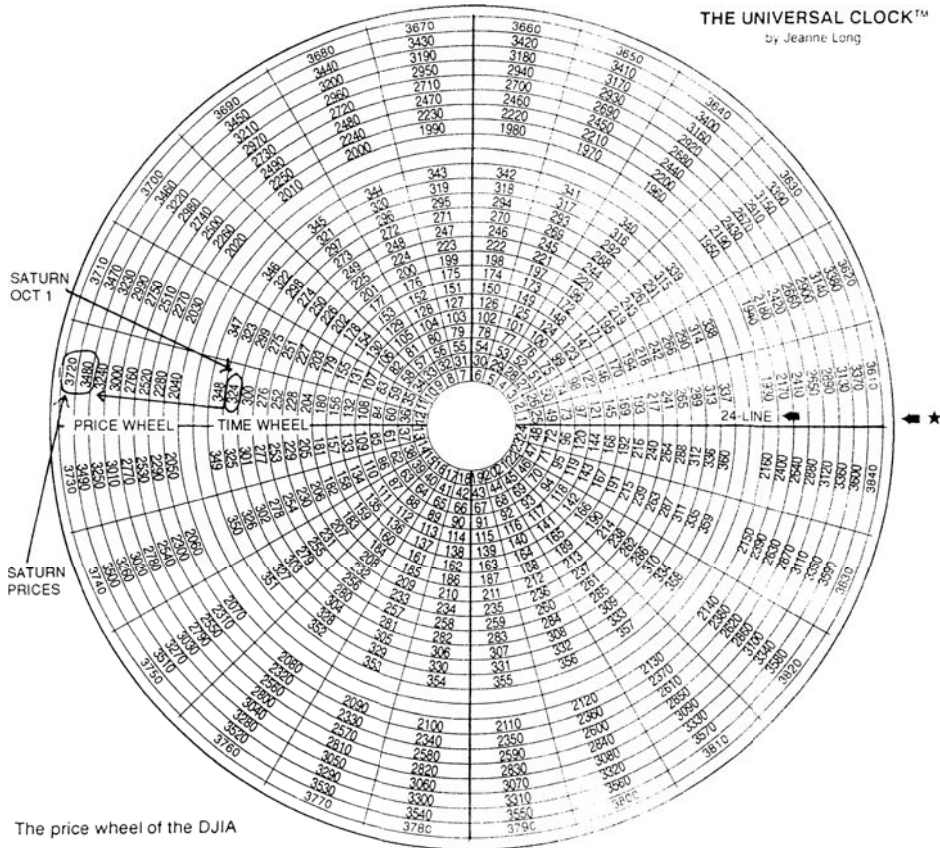
FIGURE 14.29   The DJIA Clock.

Source: Jeanne Long, *A Trader's Astrological Almanac* (1994). Reprinted with permission from PAS, Inc., e-mail: pas24@aug.com, website: www.galacticinvestor.com.

- *Conjunction*, when any two planets are on the same side of the Sun and form a straight line with the Sun.
- *Opposition*, when two planets are on opposite sides of the Sun and form a straight line with the Sun.

Examples of these events are given in Tables 14.4, 14.5, and 14.6.

The physical significance of an eclipse is said to be that it disrupts the flow of energy between two bodies by the interference of a third mass. In actuality, the gravitational forces of the two bodies are maximized when they are both on one side the Earth, as in a solar eclipse, and minimized when they are on opposite sides, as in a lunar eclipse.

In astrology, the angles between planets are viewed from the position of the Earth; this is called a *geocentric system*. When a planet lies on the line between the Earth and Sun it is in a *geocentric conjunction;* when it is aligned behind the earth it is in *geocentric opposition*. When a planet is at 90° to the earth-sun line, it is said to be *square*.[51]

---

[51] Hans Hannula, "Trading Planetary Eclipses," *Technical Analysis of Stocks & Commodities* (April 1992).

TABLE 14.4   Jupiter and Saturn Planet RX and Stations, 1996–2001

	Jupiter (Retrograde)	Jupiter (Direct)	Saturn (Retrograde)	Saturn (Direct)
1996	May 4	Sep 3	Jul 18	Dec 3
1997	Jun 9	Oct 8	Aug 1	Dec 16
1998	Jul 18	Nov 13	Aug 15	Dec 29
1999	Aug 25	Dec 20	Aug 30	
2000	Sep 29		Sep 12	Jan 12 (Jupiter not in sequence)
2001	Nov 2	Jan 25	Jan 24, Sep 26	(Saturn not in sequence)
2002		Mar 1		Feb 8

Source: Weingarten, *Investing by the Stars.*

The angles that the planets form with the Earth-Sun line are called *aspects*. For the purposes of trading, squares are considered bearish while conjunctions and oppositions are bullish; however, there are other interpretations. Conjunction and opposition may not always conform to the strict definition of forming a straight line with the Sun, but may refer instead to close proximity. Both solar and lunar eclipses are conjunctions.

Trading on Aspects

Although the clearest physical phenomena combine the sun, moon, and largest planets when they are in conjunction and opposition, there are many other combinations of planetary positions that are considered important. It may also be necessary to follow a planet from its 90° to 270° position so that you track its effect on prices through the most extreme 180° path. The calculations for lunar and solar eclipses appear at the end of this chapter, but automatically calculating the aspect of combinations of planets would be beyond the ability of even the most computer-literate traders; commercial software is available for this level of detail and can calculate the number of occurrences that an aspect corresponds to a price move or price peak of a minimum size. The simplest

TABLE 14.5   Eclipses, 1996–2001

	Lunar Eclipse	Solar Eclipse	Lunar Eclipse	Solar Eclipse
1996	Apr 4	Apr 17	Sep 27	Oct 12
1997	Mar 24	Mar 9	Sep 16	Sep 2
1998	Mar 13	Feb 26	Aug 8	Aug 22
	Sep 6			
1999	Jan 31	Feb 16	Jul 28	Aug 11
2000	Jan 21	Feb 5	Jul 16	Jul 1
		Jul 31		Dec 25
2001	Jan 9	Jun 21	Jul 5	Dec 14
	Dec 30			

Source: Weingarten, *Investing by the Stars.*

method is to select days that correspond to swing highs or swing lows and record the major aspects at that time. The larger the swing, the more significant the results. By creating a table of key reversal levels it may be possible to find those aspects that exert the greatest influence on the market. Once isolated, they may be used as a bias within a technical program.

The Moon: Buy Full, Sell New

It is known that the moon's effect on our planet is great—it is vitally connected with the movement of all fluids. The mass of the moon is about 25% that of the Earth, and larger than the planet Pluto. It is remarkably large for a moon, with a diameter of 3,476 kilometers, 27% that of the Earth. Because of its large presence and close proximity to the Earth (about 230,000 miles), the moon is also believed to affect human behavior in strange ways, especially during a new or full moon. New and full moon occurrences can be found in Table 14.6.

In an experiment conducted on an arbitrary set of futures markets for the year 1972,[52] it was shown that short-term movements of prices react with some uniformity with respect to the phases of the moon. In fact, the markets chosen for observation—silver, wheat, cattle, cocoa, and sugar—showed an uncanny ability for prices to rise following a full moon and decline after a new moon.

To find the phases of the moon, you can refer to a number of books or many calendars that note the moon phases; however, to perform any lengthy study of the moon's effect on prices, the moon's phases will need to be computerized. A program, *TSM Moon Phases*, which prints the dates of the new and full moon, is available on the Companion Website. The calculations require that dates to be converted to Julian notation, the number of days since the calendar began. At the end of the program the Julian date must be converted back to our familiar notation. The computer program includes a function to perform the conversion. An explanation of these calculations is given in the next section.

Calculation of a New Moon, Full Moon, Solar and Lunar Eclipse

Although a highly accurate calculation of planetary positions and eclipses requires consideration of many minor items, a good approximation can be found with far less difficulty.[53] To calculate the exact time of the eclipses it is first necessary to find the time of the new and full moon. The resulting times are expressed in Julian Ephemeris Days (*JDE*), also called *dynamic time* (*DT*).

---

[52]Todd Lofton, "Moonlight Sonata," *Commodities* (July 1974).

[53]Jean Meeus, *Astronomical Algorithms* (Willmann-Bell, Inc., P.O. Box 35025, Richmond, VA 23235), Chapter 52. A shorter version appears in another volume, *Astronomical Formulae for Calculators*. For complete accuracy, the author refers readers to the *Astronomical Almanac*, or the *Canon* by Mucke and Meeus. The material in this section combines both versions of Meeus' work.

TABLE 14.6 New and Full Moons, Eclipses 1996–2001

Year	New	Full	New	Full	New	Full	New	Full	New	Full	New	Full
1996	Jun 16	Jan 5	Jan 20	Feb 4	Feb 18	Mar 5	Mar 19	Apr 4	Apr 17	May 3	May 17	Jun 1
	Dec 10	Jul 1	Jul 15	Jul 30	Aug 14	Aug 28	Sep 12	Sep 27	Oct 12	Oct 26	Nov 11	Nov 25
		Dec 24										
1997	Jan 9	Jan 23	Feb 7	Feb 22	Mar 9	Mar 24	Apr 7	Apr 22	May 6	May 22	Jun 5	Jun 20
	Jul 4	Jul 20	Aug 3	Aug 18	Sep 1	Sep 16	Oct 1	Oct 16	Nov 14	Nov 30	Dec 14	Dec 19
1998	Jun 24	Jan 12	Jan 28	Feb 11	Feb 26	Mar 13	Mar 28	Apr 11	Apr 26	May 11	May 25	Jun 10
	Dec 18	Jul 9	Jul 23	Aug 8	Aug 22	Sep 9	Sep 20	Oct 5	Oct 20	Nov 4	Nov 19	Dec 3
1999	Jun 13	Jan 2	Jan 17	Jan 31	Feb 16	Mar 2	Mar 17	Mar 31	Apr 16	Apr 30	May 15	May 30
	Dec 22	Jun 28	Jul 13	Jul 28	Aug 11	Aug 26	Sep 9	Sep 25	Oct 9	Oct 24	Nov 23	Dec 7
2000	Jan 6	Jan 21	Feb 5	Feb 19	Mar 6	Mar 20	Apr 4	Apr 18	May 4	May 18	Jun 2	Jun 16
	Jul 1	Jul 31	Aug 15	Aug 29	Sep 13	Sep 27	Oct 13	Oct 27	Nov 11	Nov 25	Dec 11	Dec 25
2001	Jun 21	Jan 9	Jan 24	Feb 8	Feb 23	Mar 9	Mar 25	Apr 8	Apr 24	May 7	May 23	Jun 6
	Dec 14	Jul 5	Jul 20	Aug 4	Aug 19	Sep 2	Sep 17	Oct 2	Oct 16	Nov 1	Nov 15	Nov 30
		Dec 30										

Source: Weingarten, *Investing by the Stars.*

Lunar Phases

The new and full moons are simply two of the four lunar phases, measuring when the excess of the apparent longitude of the moon over the apparent longitude of the Sun is 0°, 90°, 180°, and 270°. The time of these phases are given by the Julian date:

1. $JDE = 2451\,550.09765 + 29.530588\,853\,k + 0.0001337\,T^2$
   $-\,0.000000\,150\,T^3 + 0.000000000\,73\,T^4$

   where an integer value of k gives the new moon (i.e., 1, 2, 3, . . .) and an integer value increased by 0.25, 0.50, or 0.75 (e.g., 1.25, 1.50, 1.75) gives the three quarter positions of the moon, respectively. The value $k = 0$ corresponds to the new moon of January 6, 2000. Negative values indicate times prior to the year 2000. T is the time in Julian centuries since 2000, found by $T = k/1236.85$.

   In order to check your results, k is approximately equal to $(year - 2000) \times 12.3685$ and year is expressed in decimal notation (e.g., the end of March 1997 is 1997.25). Using this formula, we can get a list of dates for the full moon from about 1995 through the year 2000 (approximately 62) by creating an array called *JDE* and finding the 62 dates of the new and full moons ending at the year 2000 using the program *TSMMoonPhases* on the Companion Website.

   For each value of k in the program, corrections must be made to find the exact time of maximum solar or lunar eclipse using the following values:

2. Eccentricity of the Earth's orbit around the Sun:

$$E = 1 - 0.002516T - 0.0000074T^2$$

3. The Sun's mean anomaly at time *JDE:*

$$M = 2.5534 + 29.10535669k - 0.0000218T^2 - 0.00000011T^3$$

4. The moon's mean anomaly:

$$M' = 201.5643 + 385.816935\,28\,k + 0.0107438T^2 + 0.00001239T^3 - 0.000000058T^4$$

5. The moon's argument of latitude:

$$F = 160.7108 + 390.67050274k - 0.0016341T^2 - 0.00000227T^3 + 0.000000011T^4$$

6. The ascending node of the lunar orbit:

$$\Omega(Omega) = 124.7746 - 1.56375580k + 0.0020691T^2 + 0.00000215T^3$$

7. The eccentricity of the Earth's orbit around the Sun:

$$E = 1 - 0.002516T - 0.0000074T^2$$

8. The following calculation, based on F and Q, gives the basis for the solar or lunar eclipse. If F differs from 0°, 180° or 360° by less than 13.9°, then an eclipse is certain. If F is more

than $21°$ from these phases, there is no eclipse. If F falls between these values then there is no eclipse if $|\sin F| > 0.36$. The next components needed for corrections are:

$$F_1 = F - 0°.02665 \sin \Omega$$

$$A_1 = 299°.77 + 0°.107408\, k - 0.009173T^2$$

9. Then to find the exact time of the full solar (or lunar) eclipse, the following corrections (in days) should be added to the time of the mean conjunction given by *JDE* in the first formula above (smaller quantities have been omitted):

Time of maximum eclipse = JDE + (*lunar* or *solar* component below) + 0.0161 sin $2M'$ − 0.0097 sin $2F_1$ + 0.0073 × E × sin $(M' - M)$ − 0.0050 × E × sin $(M' + M)$ − 0.0023 sin $(M' - 2F_1)$ + 0.0021 × E × sin $2M$ + 0.0012 sin $(M' + 2F_1)$ + 0.0006 × E × sin $(2M' + M)$ − 0.0004 sin $3M'$ − 0.0003 × E × sin $(M + 2F_1)$ + 0.0003 sin A_1 − 0.0002 × E × sin $(M - 2F_1)$ − 0.0002 × E × sin $(2M' - M)$ − 0.0002 Ω

solar component = −0.4075 sin M' + 0.1721 × E × sin M

lunar component = −0.4065 sin M' + 0.1727 × E × sin M

10. $P = 0.2070 \times E \times \sin M + 0.0024 \times E \times \sin 2M - 0.0392 \sin M' + 0.0116 \sin 2M'$
          $- 0.0073 \times E \times \sin (M' + M) - 0.0067 \times E \times \sin (M' - M) + 0.0118 \sin 2F_1$

11. $Q = +5.2207 - 0.0048 \times E \times \cos M + 0.0020 \times E \times \cos 2M - 0.3399 \cos M'$
          $- 0.0060 \times E \times \cos (M' + M) + 0.0041 \times E \times \cos (M' - M)$

12. $W = |\cos F_1|$

13. $\gamma\,(\text{gamma}) = (P \cos F_1 + Q \sin F_1) \times (1 - 0.0048W)$

14. $u = 0.0059 + 0.0046 \cos M - 0.0182 \cos M' + 0.0004 \cos 2M' - 0.0005 \cos (M + M')$

Solar Eclipses

For a solar eclipse, γ represents the shortest distance from the axis of the moon's shadow to the center of the Earth in units of the equatorial radius of the Earth (the distance from the center to the surface of the Earth at the equator). Its value is positive if the axis of the shadow is passing north of the Earth's center, and negative if it is passing south. When γ is less than +0.9972 and greater than −0.9972 the solar eclipse is *central*, that is, there is a line of central eclipse on the surface of the Earth (see Figure 14.30).
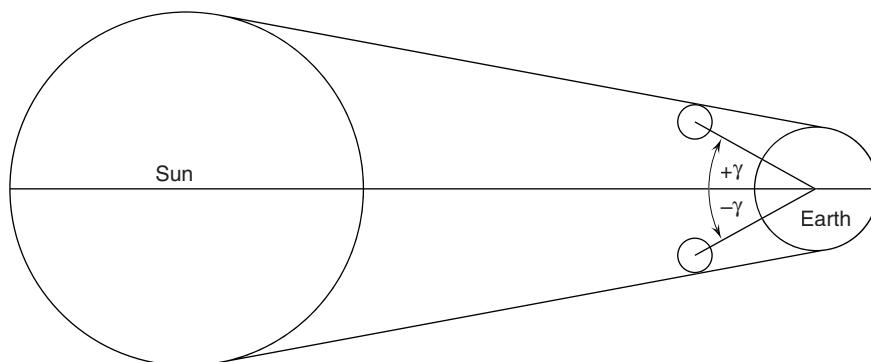

FIGURE 14.30   Geometry of a solar eclipse.

The value u gives the radius of the moon's *umbral cone* on the fundamental plane, which passes through the center of the Earth perpendicular to the axis of the moon's shadow (in units of Earth's equatorial radius). The radius of the *penumbral cone* in the fundamental plane is $u + 0.5460$. Based on these values, the following situations exist:

$\|\gamma\|$ is between 0.9972 and $1.5432 + u$	The eclipse is not central
$\|\gamma\|$ is between 0.9972 and 1.0260	Part of the eclipse may touch the polar regions
$0.9972 < \|\gamma\| < 0.9972 + \|u\|$	The combination of a noncentral total or annular eclipse
$\|\gamma\| > 1.5432 + u$	No eclipse if visible from the Earth's surface
For a central eclipse, $u < 0$	The eclipse is total
$u > 0.0047$	The eclipse is annular
$0 < u < 0.0047$	The eclipse is either annular or annular-total.

To remove the ambiguity of this last situation, calculate

$$\omega = 0.00464 \cos W, \text{ where } \sin W = \gamma$$

If $u < \omega$ then the eclipse is annular-total; otherwise it is annular.

The greatest magnitude for a partial solar eclipse is reached at the point on the Earth's surface that comes closest to the axis of the shadow at

$$\textit{Greatest magnitude of a solar eclipse} = \frac{1.5433 + u - \|\gamma\|}{0.5461 + 2u}$$

Lunar Eclipses

For a lunar eclipse, γ is the least distance for the center of the moon to the Earth's shadow (see Figure 14.31). The value γ is positive if the moon's center is passing north of the axis of the shadow, and negative if it is passing south. At the distance of the moon, the penumbra radius $\rho = 1.2847 + u$ and the umbra radius $\sigma = 0.7404 - u$. The magnitude of the lunar eclipse is

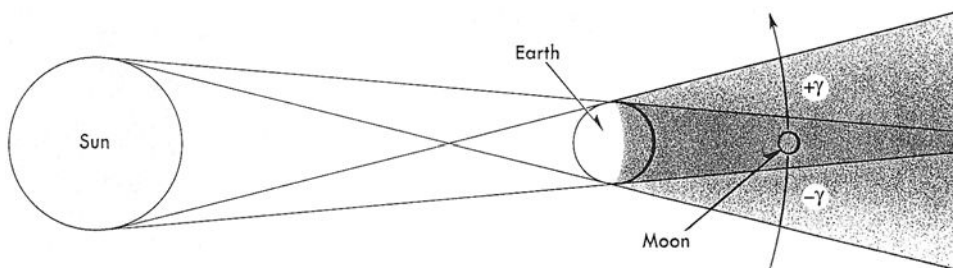$$\text{penumbral eclipses: } (1.5572 + u - \|\gamma\|)/0.5450$$

$$\text{umbral eclipses: } (1.0129 - u - \|\gamma\|)/0.5450$$

If the magnitude is less than zero, there is no eclipse. The *semidurations* of the partial and total phases in the *umbra* are calculated as

$$P = 1.0129 - u$$
$$T = 0.4679 - u$$
$$n = 0.5458 + 0.0400 \cos M'$$

FIGURE 14.31   Geometry of a lunar eclipse.

and the semidurations in minutes are

$$Partial\ phase = \frac{60}{n}\sqrt{P^2 - \gamma^2}$$

$$Total\ phase = \frac{60}{n}\sqrt{T^2 - \gamma^2}$$

Example: Solar Eclipse of May 21, 1993

The following values can be used to verify your calculations:
   May 21 is the 141st day of the year, therefore

$$k = 1993 + 141/365 = 1993.38$$

$$T = 1993.38/1236.85 = 1.6116$$

$$JDE = 2449128.5894$$

$M = 135°.9142$	$P = 0.1842$
$M' = 244°.5757$	$Q = 5.3589$
$F = 165°.7296$	$\gamma = 1.1348$
$\Omega = 253°.002$	$u = 0.0097$
$F_1 = 165°.7551$	

Because $180° - F$ is between $13°.9$ and $12°.0$ the eclipse is uncertain, and because $|\gamma|$ is between 0.9972 and $1.5433 + u = 1.553$, the eclipse is partial. By calculating the greatest magnitude of a solar eclipse, we get 0.740. Because F is near $180°$, the eclipse occurs near the moon's descending node, and because γ is positive, the eclipse is visible in the northern hemisphere of the Earth. By adding the corrections to *JDE*, the final time of maximum eclipse is 2449129.0979, which corresponds to May 21, 1993, at 14h21m0s TD. This differs from the exact value of 14h20m14s TD by less than 1 minute.

Software

As you would expect, the past 10 years has seen software of every form come into the public domain. By searching the Internet for "Financial Astrology Software" you can find services that calculate various astronomical and astrological events and interpret them. However, in order to use this method for trading it is best that some history of performance be tested. To do that, you will need to produce the numbers in a way that can be incorporated into an automated trading strategy. Without doing that, we have a tendency to remember particularly good trades and forget the more frequent small losing ones.

The U.S. Government provides the most definitive calculator for eclipses and other events at http://eclipse.gsfc.nasa.gov/JSEX/JSEX-index.html. Rather than program the formulas on a strategy development platform, you can calculate a number of events forward and backward and input those dates for testing.

Pattern Recognition

P attern recognition forms the basis for most trading systems. Patterns are most obvious in traditional charting, which is primarily the identification of common formations; even moving averages attempt to isolate, using mathematical methods, what has been visually determined to be a trend. Traders have always looked for patterns in price movement. Because the earliest technicians were not equipped with computers, their conclusions are considered market lore rather than fact, and are handed down from generation to generation as proverbs, such as "Up on Monday, down on Tuesday," "Locals even-up on Fridays," and "Watch for key reversals." Because these three sayings have endured, they are candidates for analysis later in this chapter. Readers should also note the early work of Arthur Merrill, whose well-known pattern studies are still quoted; these studies are referenced throughout this chapter.[1]

The earliest technical systems based on patterns were of the form: "If after a sharp rise the market fails to advance for 3 days, then sell." As computers became more powerful, more complex approaches could be taken. For example, by observing the closing prices starting on an arbitrary day, *all* patterns of higher and lower closes can be recorded to find their tendency to repeat. A computer is well-equipped to perform this task. First, the 2-day combinations of price changes—*up-up*, *up-down*, *down-up*, and *down-down*—are tallied to see if there is a greater chance of, for example, an up day following an up day and a down day following a down day. Next, the eight 3-day patterns are tested for the same possibility of a high-probability price move, then sixteen 4-day patterns, and so on. When you get to larger sequences, such as *up-up-down-up-down-up*, every six consecutive prices must be tested for that pattern. From the price moves that follow this pattern, we can conclude whether it has any predictive ability. This approach as well as the combination of events used to forecast profitable situations are discussed later in this chapter.

Pattern recognition may appear to be more of a game than a business, but it is a source of many valuable ideas as well as false paths. Figure 15.1, a graph of the New

---

[1] Arthur A. Merrill, *Behavior of Prices on Wall Street* (Chappaqua, NY: Analysis Press, 1966).
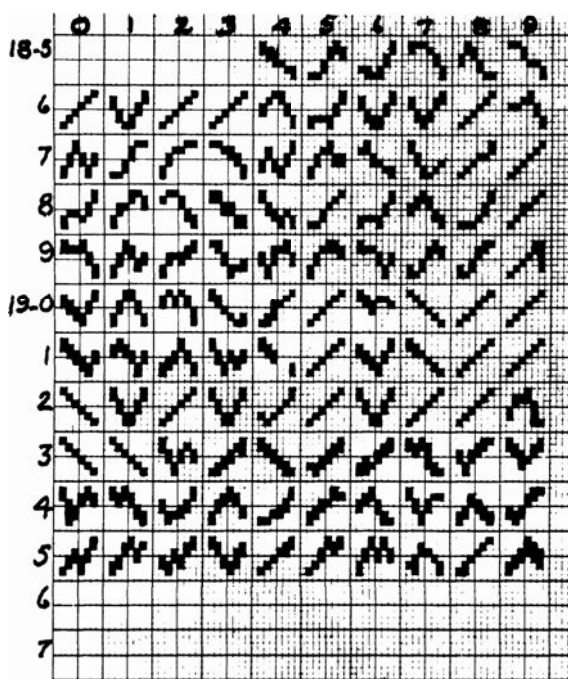
FIGURE 15.1   Graph of the New York stock market.

York Stock Exchange (1854–1959), shows a very simple way to visualize the price pattern throughout the year. Each box represents the net movement of consecutive 2-month intervals, and is either up or down. When viewing all the patterns for 104 years in one glance, it is difficult not to count the recurrences of the more obvious patterns and then look for the formations that precede them in order to see whether they could be predictive. For example, in 1922, 1924, and 1927, there were sharp advances in the market; the years preceding those showed identical *V* patterns. When we look for all other *V* patterns, we see that 3 of 5 were followed by the same bull moves. Another pattern that stands out is that of two consecutive years of sharp rise, 1862–1863, 1908–1909, 1918–1919, and 1927–1928; in no case was there a third consecutive year but neither the preceding nor following years seem consistent.

Patterns frequently provide the foundation for a trading method or the motivation to begin the development of a method. They have been applied in many ways to price analysis, from the time of day to place an order to the complex relationship of price, volume, and open interest. Daily trading opportunities may be a function of patterns based on the strength or weakness of the daily opening price. These are discussed in the next section. Weekly and weekend price patterns are studied, as well as types of reversals and their effects. These techniques can be used sequentially by following a weekday pattern with a weekend pattern, or they can confirm the results of each other when used together. The end of this chapter discusses more general issues in pattern recognition.

Because patterns are implicit in many trading systems throughout this book, readers are encouraged to use the index if they are looking for a discussion of a specific pattern. For example, intraday volume patterns will be found in the next chapter.

PROJECTING DAILY HIGHS AND LOWS

Statisticians claim that the best forecast of tomorrow's price is today's price; for longer-term projections, the best forecast is the mean. They state that no one has been able to reliably account for the numerous combinations of price moves that would be needed to accurately project the changes that result in the high and low prices for tomorrow. Nevertheless, many traders project these values and seem to use them with success.

Pivot Points

The simplest approach to projecting tomorrow's highs and lows is to base those figures on the average price of today plus or minus a value that somehow relates to the current trading range, or volatility. One technique, based on a *pivot point*,[2] projects two levels of support and resistance:

$$
\begin{aligned}
\text{Pivot point} \quad & P = (H + L + C)/3 \\
\text{Resistance 1} \quad & R1 = 2 \times P - L \\
\text{Support 1} \quad & S1 = 2 \times P - H \\
\text{Resistance 2} \quad & R2 = (P - S1) + R1 \\
\text{Support 2} \quad & S2 = P - (R1 - S1)
\end{aligned}
$$

As an example, take the high, low, and close of the day to be 12.50, 11.50, and 11.75. Then

$$
\begin{aligned}
P &= (12.50 + 11.50 + 11.75)/3 = 11.916 \\
R1 &= 2 \times 11.916 - 11.50 = 12.332 \\
S1 &= 2 \times 11.916 - 12.50 = 11.332 \\
R2 &= (11.916 - 11.332) + 12.332 = 12.914 \\
S2 &= 11.916 - (12.332 - 11.332) = 10.918
\end{aligned}
$$

The projected support and resistance levels show a downward bias consistent with the close that was in the lower part of the trading range.

Another method takes the average range (*high − low*) over the past 14 days, then adds and subtracts ½ of that range to yesterday's close, so that the projection is symmetric around the prior close and equal to the average range.

---

[2]Mark Etzkorn, "All in a Day's Work," *Futures* (January 1995). This technique is attributed to William Greenspan.

DeMark's Projected Ranges

Another author and trader, Tom DeMark,[3] has based his projections on the positioning of today's opening price relative to yesterday's closing price, which gives him added timely information. If today's open is higher than the previous close, the projections are biased upwards; if lower than the close, they are biased downwards.

If today's close is below today's open, then

Tomorrow's projected high = $(H + C + 2 \times L)/2 - L$
Tomorrow's projected low = $(H + C + 2 \times L)/2 - H$

If today's close is above today's open, then

Tomorrow's projected high = $(2 \times H + L + C)/2 - L$
Tomorrow's projected low = $(2 \times H + L + C)/2 - H$

If today's close is the same as today's open then

Tomorrow's projected high = $(H + L + 2 \times C)/2 - L$
Tomorrow's projected low = $(H + L + 2 \times C)/2 - H$

Once the basic formula is determined, which biases tomorrow's projection in the direction of today's close relative to today's open, the *projected high* is found by removing two units of the low price, and the *projected low* is found by removing two units of the high price from the formula. This effectively shifts the projection in the direction of the new high by one-half the difference of today's close and today's low, $(C - L)/2$, and the new low by one-half the difference of today's close and today's high, $(H - C)/2$.

Comparing the Two Ranges

Because the two methods of calculating the projected ranges use the same basic prices, combining them with slightly different arithmetic, it is not clear whether they give values that are the same or different. For example, if the Swiss franc had a high, low, and close of 6600, 6450, and 6500, respectively, the first pivot method would give a projected first resistance level (high) of $2 \times (6600 + 6450 + 6500)/3 - 6450 = 6583$. DeMark's projected high is different, based on the way today's opening price relates to the previous close. If the open is lower than the previous close, we get $(6600 + 6500 + 2 \times 6450)/2 - 6450 = 6550$; if the open is higher, we have $(2 \times 6600 + 6450 + 6500)/2 - 6450 = 6625$; and, if the open is about the same as the previous close, then the projected high is $(6600 + 6450 + 2 \times 6500)/2 - 6450 = 6575$. It seems reasonable that the pivot point method returns a value close to DeMark's neutral case when the market opens unchanged from the previous close.

---

[3]Tom DeMark, *The New Science of Technical Analysis* (New York: John Wiley & Sons, 1994).

TIME OF DAY

Market participants, especially floor traders and day traders, are the cause of periodic movement during the day. Angas called these the "tides of the daily prices." Over the years, the great increase in participants has added liquidity to nearly all markets but has not altered the intraday time patterns. Even high frequency trading seems to have added huge trading volume without altering the basic patterns. The classic pattern in both stocks and futures shows the greatest volume near the open, the next highest volume near the close, and the lowest volume at midday.

Trading Habits Form Lasting Patterns

There are a number of reasons for this normal pattern of trading. Many investors evaluate their positions and study market reports in the evenings and then place their orders in the morning. In Europe, where business begins six hours earlier, economic reports are released before the main trading sessions in the United States, and can affect the early market direction. However, U.S. business reacts to U.S. reports, so U.S. markets readily change direction once U.S. business begins. Although futures in all the major index markets are trading throughout the night, price movements during these off hours are not considered important to the direction of prices the next day. Even now, European markets are open until the U.S. stock markets close; however, that's 10 P.M. in Frankfurt, and most traders have gone home. Only a few brokerage houses have desks that will accept orders late in the evening. The only practical way of trading is using an electronic order entry program. With all the globalization, overnight trading has not materially changed the patterns that occur during the day sessions.

The close of the day is a common time for trading because many investors believe that the closing price best represents the correct value. In the futures markets, a large part of the daily volume is the result of day trades by floor traders, or *locals.* Positions entered in the morning will be closed out by the end of the day to avoid depositing the margin required of positions held overnight. Accounts are settled using the closing price.

Once the opening orders have been executed—the order flow that has accumulated from decisions made since the last close—volume declines. Scalpers and floor traders are active during the opening minutes but frequently have a mid-morning coffee break, and more take a lunch break; this natural phenomenon causes liquidity to decline and may result in a temporary price reversal. All traders and investors develop habits of trading at particular times. Some prefer the opening, others 10 minutes after the open. Large funds and managed accounts have a specific procedure for entering the market, such as using close-only orders. Institutions with large orders must trade during times when the volume is heavy in order to keep slippage under control, or they may take advantage of VWAP orders (volume-weighted average price) where partial orders are automatically

fed into the market based on an algorithm that balances time intervals and volume.[4] Without these orders, most institutions will favor trading at the more active times.

A day trader must watch certain key times. The opening moments of the trading session are normally used to assess the situation. The stock market will gap up or down to align itself with the futures market, which has been trading continuously. Most other markets have electronic sessions in which the prices reflect current news. The day session opens with the bid-asked at the level already trading in the electronic session; therefore, there is little opportunity for profit in the first few seconds of trading. After a short delay, as the first burst of activity slows, floor trader will sell a strong open and buy a weak one; this means the trade must be *evened up* later, reinforcing the opening direction. On a strong open without a downward reaction, all local selling is absorbed by the market, and later attempts of the locals to liquidate will hold prices up, preventing a reversal. Floor traders can normally be expected to take the opposite position to the opening direction, usually causing a reversal early in the session.

Tubbs' Intraday Patterns

In Chapter 13 of Tubbs' *Stock Market Correspondence Lessons* ("Tape Reading") he explains the six dominant patterns in the stock market (based on a 10:00 A.M. to 3:00 P.M. session).

1. If a rally after the open has returned to the opening price by 1:00 P.M., the day is expected to close weaker.

2. If the market is strong from 11:00 A.M. to 12:00 P.M., it will continue from 12:00 P.M. to 1:00 P.M.

3. If a reversal from 1:00 P.M. to 1:30 P.M. finds support at 1:30 P.M., it will close strong.

4. If the market has been bullish until 2:00 P.M., it will probably continue until the close and into the next day.

5. A rally that continues for 2 or 3 days (as in Pattern 4) will most likely end on an 11:00 A.M. reversal.

6. In general, a late afternoon reaction down after a strong day shows a pending reversal.

Putting these together, the following two patterns (among others) can be expected:

1. A strong open with a reversal at 11:00 A.M. not reaching the opening price, then strength from 11:00 A.M. to 1:00 P.M., a short reversal until 1:30 P.M., and then a strong close; according to Pattern 5, there will be a strong open the following day.

2. A strong open that reverses by 11:00 A.M., continuing lower until 1:00 P.M., reverses again until 1:30 P.M., will close weak.

---

[4]At the time of this writing, source code for automatically generating VWAP orders can be found on www.codeproject.com/KB/recipes/VWAP.aspx.

TABLE 15.1	Merrill's Hourly Stock Market Patterns					
	Time during Trading Session					
	10:00	**11:00**	**12:00**	**1:00**	**2:00**	**3:00**
1962	–	+	–	–	+	–
1963	+	+	+	+	+	–
1964	+	+	–	–	+	–
1965	+	+	–	–	–	–

Merrill's Intraday Patterns

Merrill's work shows the hourly pattern of the stock market in Table 15.1, where a plus or minus sign indicates rising or falling prices during the preceding period. During the four years shown, the grid clearly indicates a strong opening and follow-through to 11:00 A.M. During 1963, the bullishness lasted all day, trailing off at the close. Except for 1963, we cannot say from this table that there was a bullish or bearish price move because there is no volatility, or average price move, associated with these hourly intervals.

The pattern in Table 15.1 supports what we would expect as a normal intraday pattern. Prices begin strong and continue the first hour. Having exhausted the early orders, volume decreases, and price reverses due to lack of buyers. By 2 P.M. the short sellers who faded the first-hour move cover their positions, causing a move back in the direction of the opening hour. The only inconsistency in the pattern is the weak close, which contradicts the strong open. Had there been both a strong open and a strong close, we could declare these as bull years. Nevertheless, the pattern of a weaker close persisted for at least four years, offering an opportunity for day traders.

Updating Intraday Time Patterns, 2000–2011

Every decade is unique, and 2000 was no exception. Prices began at extreme highs, fell for three years, languished for a year, then rallied only to collapse in 2008 and rally again. We can expect a bias to the upside during a sustained bullish period, such as the 1990s, but the last decade is much more interesting. It gives us an opportunity to understand the intraday patterns under very different market scenarios. The 11 years in Figure 15.2 can be separated into four distinct periods:

1. 2001–2002, strong bear market
2. 2003–2007, normal bull market
3. 2008, extreme collapse
4. 2009–May 2011, strong bull market
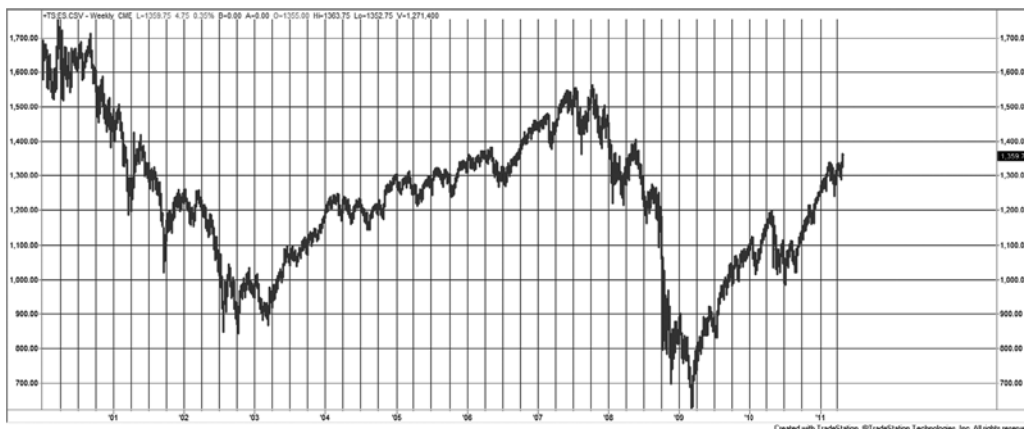
FIGURE 15.2 S&P back-adjusted futures from 2000 through May 2011.

First using the entire 11 years, the patterns are evaluated in three ways:

1. The cumulative change in returns for each 30-minute period throughout the day.
2. The change relative to the direction of the open from the previous close (as in Merrill's Table 15.1).
3. The change relative to the opening 30 minutes of trading.

These are shown in Figure 15.3 as *Average move*, *Same as previous close*, and *Same as first bar*. The bottom line, *Average move*, shows that the net returns over the full period were near zero. The top line, *Same as previous close*, indicates that there was a modest continuation of the prior close to end of first bar direction throughout the day. That is, if the end of the first bar was higher than the previous close, then it was likely to close even higher, and if it opened lower, then it would continue lower throughout the day. The middle line, *Same as first bar*, shows that there is essentially no change after the opening bar, so that most of the daily move occurs overnight and on the opening bar.
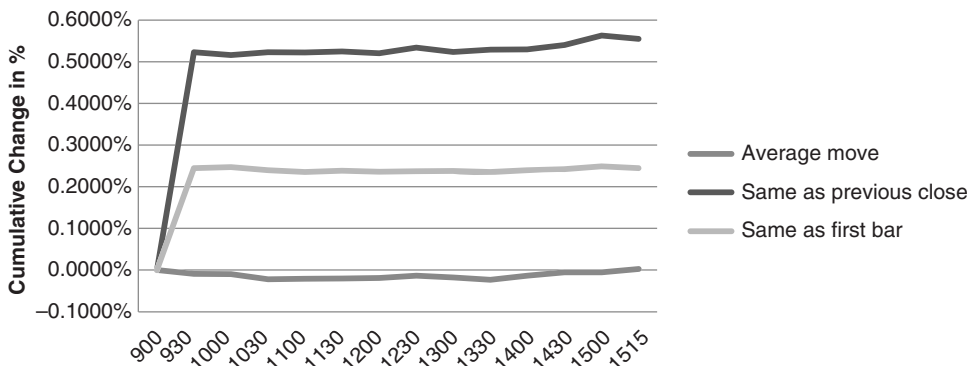

FIGURE 15.3 S&P futures intraday time patterns for 2000 through May 2011.
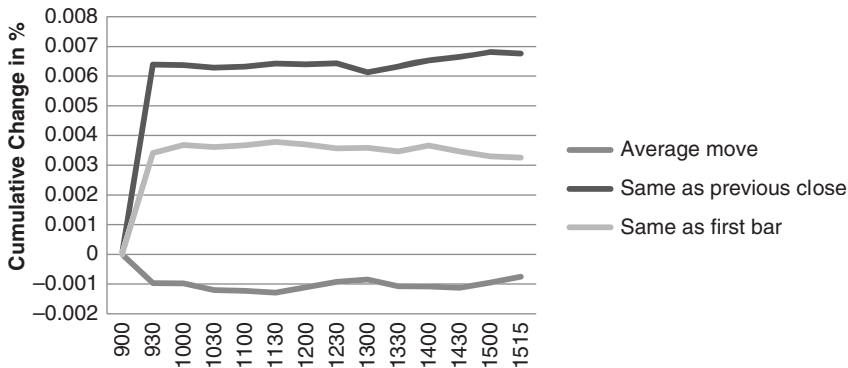
FIGURE 15.4a S&P 2001–2002, strong bear market.

If the 11 years are separated into the four periods that represent clear regimes, the results are more interesting. Figures 15.4a–d show the results, which can be summarized as

a. 2001–2002, a strong bear market shows that the average move ends the day lower and the direction at the close of the first bar persists slightly throughout the day.

b. 2003–2007, a normal bull market shows the average move steadily increasing throughout the day as did the direction from the previous close. However, prices tended to reverse slightly relative to the direction of the opening bar.

c. 2008, the extreme collapse due to the subprime crisis, is the most consistent. All three measures continued in the direction of the opening moves and the overall market direction.

d. 2009–2011, a strong bull market, shows in the average move, but is different in the other measurements. Rather than continuing in the direction given by the prior close to end of first bar, or from the open to the end of the first bar, prices reverse, giving back earlier gains or losses.
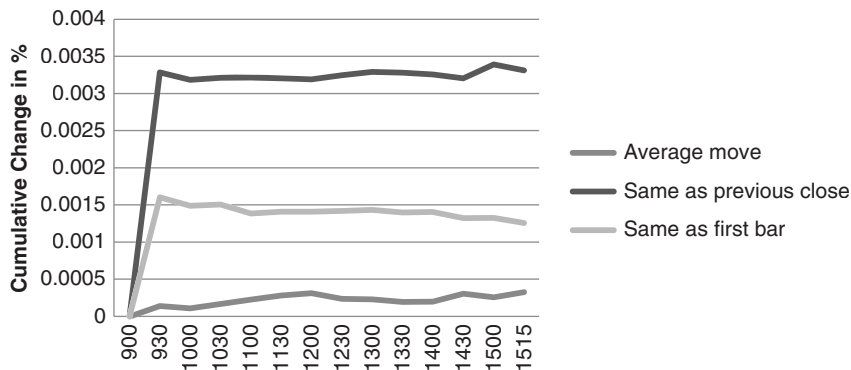
**FIGURE 15.4b** S&P 2003–2007, normal bull market.
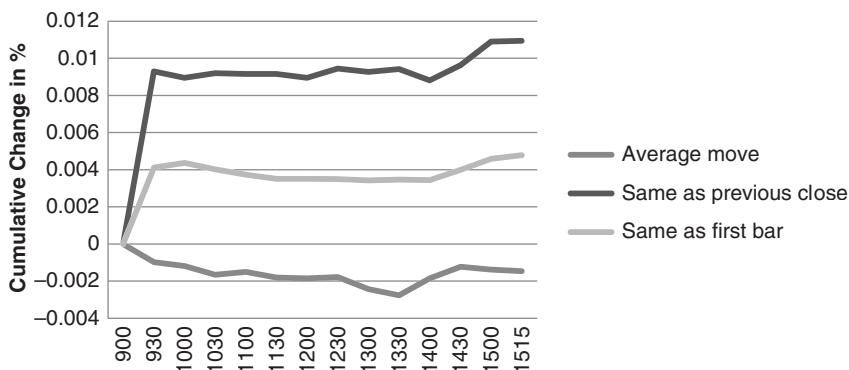
FIGURE 15.4c   S&P 2008, extreme collapse.

Overall, we can say that there is a small profit that can be extracted by taking a position at the end of the first bar in the direction of the previous close, and holding to end of first bar. Unfortunately, the four cases were classified after the fact; therefore, it will take more time to recognize when a bull or bear market is in effect. The only case that fails is the strong bull market, which still leaves reasonable opportunity in the other three scenarios.

The overall intraday time pattern of the S&P can be summarized in the same way as Merrill in the following table, where a "+" represents a net gain from the close of the previous bar and "−" a net loss:

Time	1000	1100	1200	100	200	300	400
2000	−	+	+	+	+	+	+
2008	−	−	−	−	+	+	+

For the full 11-year period, the opening bar was generally lower, but each subsequent bar was higher. From 2008 through May 2011 the first half of the day was lower and the last half higher. This does not show whether the day was a net gain or loss, only the relative direction of prices.
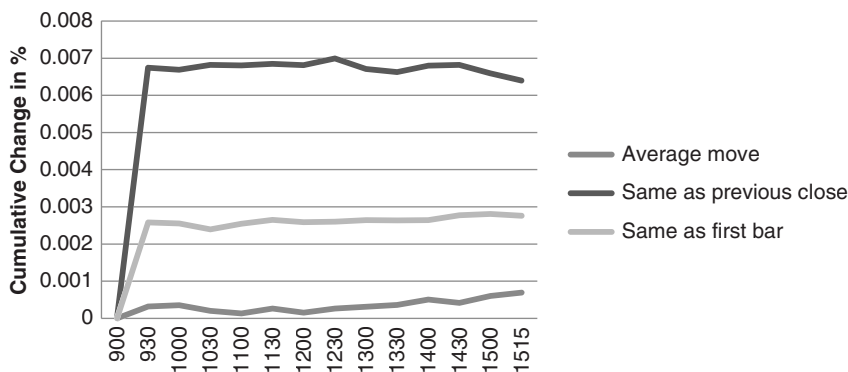
FIGURE 15.4d   S&P 2009–May 2011, strong bull market.

EURUSD Time Patterns

Is the pattern of the S&P typical of other markets? We would expect it to be similar to other index markets, but other sectors, such as commodities and currencies, could have unique patterns of their own. As an example, the same analysis was run on the EURUSD currency futures for the same 11-year period from 2000 through May 2011. Results are shown in Figure 15.5.

The overall pattern is only slightly different from the S&P. There was a net gain in the EURUSD, shown by the *Average move*; however, that gain was all in the overnight move. From the end of the first bar, which occurs at 3 A.M. in New York (9 A.M. in Frankfurt), prices tend to reverse. The move relative to the direction of the open-close of the first bar is essentially flat. Another way of viewing these patterns is by indicating which markets are open during each time interval. In the following table, trading begins at 2 A.M. in Europe and ends at 3 P.M. in New York. The first 5 hours are actively traded in Europe but not the United States. There are 3 hours of overlapping trading after which Europe closes (at 4 P.M. local time). The U.S. trades for another 5 hours, with only small volume coming from Europe.

	Europe is open at 200					U.S. is open			Europe is closed				
Time	300	400	500	600	700	800	900	1000	1100	1200	1300	1400	1500
2000	−	+	−	−	−	+	+	−	+	+	+	+	+
2008	+	+	−	−	−	+	−	−	+	+	+	+	+

The table shows that, from 2000 through May 2011, despite the net gain in the EURUSD, that gain came mostly when the U.S. session was open. During the past 2½ years, the larger gain in the EURUSD is reflected in both the European open and in the U.S. session. It appears as though the U.S. trading is driving the U.S. dollar lower.
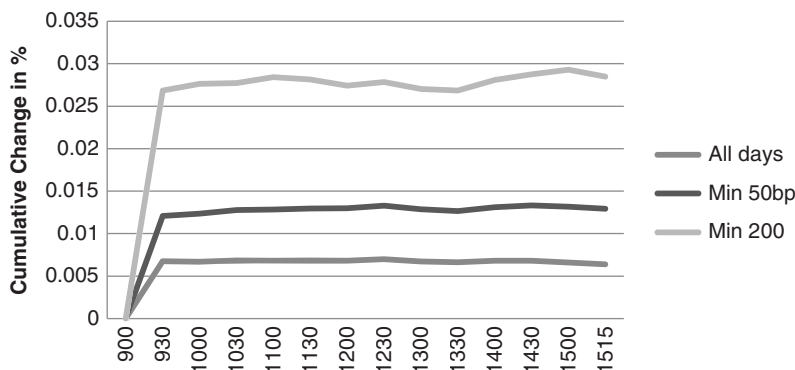
Refining the Patterns for Trading

To increase the chance of a profitable trade, it might be best to observe only stronger or weaker openings. For example, if the end of the first 30 minutes of S&P trading shows a gain greater than 1% from the previous close, we can expect the entire day to be more
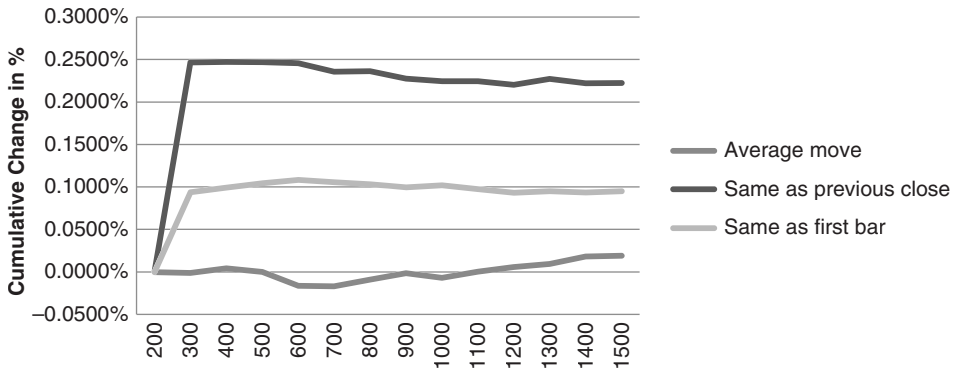
FIGURE 15.5 Time patterns for EURUSD futures from 2000 through May 2011.

FIGURE 15.6 S&P futures time patterns, 2009–2011,with 50 bp and 200 bp minimum gap filters.

volatile than a day that posted a gain of only 5 or 10 basis points. Restricting trading to days of higher volatility will reduce the trading opportunities but should improve results. By allowing very small opening price changes, we include days that might drift sideways because of the lack of news.

Repeating the same test of the S&P for 2009 through May 2011, but with the addition of a 50 basis point filter, Figure 15.6 shows that the average gap from the previous close to the end of the first 30-minute bar nearly doubles to 1.2% with a 50 bp threshold, and the balance of the day continues modestly in the same direction as the gap, while the unfiltered test reverses slightly. If the minimum gap filter is increased to 200 bp, the balance of the days seems to continue at perhaps a slightly increased rate, but certainly the same pattern. As the gap filter increases, it is possible that prices will show a reversal pattern.

One possibility for trading is to first decide that price will continue to move in the direction of the opening gap, then wait for a pullback. Even in Figure 15.6, which shows the average of a large number of days, there is a pullback low at 13:30 and a rally shortly afterwards. Individual days will have a much wider price range and higher volatility that could produce profits.

A Program for Time Patterns

Readers who would like to pursue the study of intraday time patterns will find a program, *TSM Intraday Time Patterns*, on the Companion Website. It creates a table of returns for each time bar that includes the average change, the change relative to the close to end of first bar direction, and the change relative to the open to close of the first bar. Price changes can be specified as differences or returns, and an opening gap filter can be used, specified as a percentage return.

Intraday Highs and Lows

In selecting a place to enter the market for a single day trade, it would be a great advantage to know the time of day at which the highest or lowest price is likely to occur. It
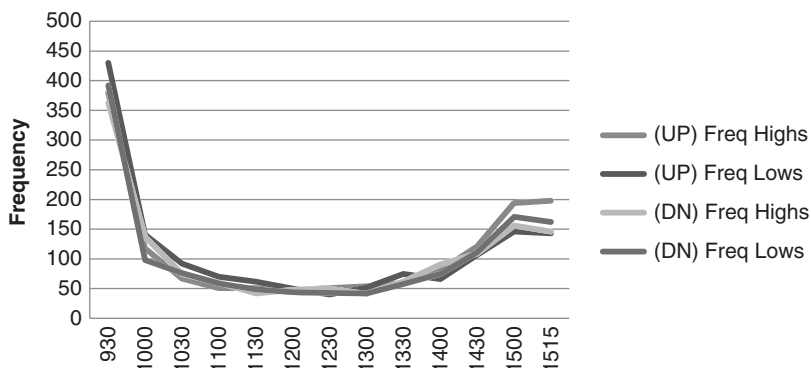
FIGURE 15.7 Frequency of intraday highs and lows for SP futures, 2000 through May 2011.

would also be useful to know if that pattern changes when the market opens higher or lower, and when there is a large price gap on the open. The same markets, the S&P and the EURUSD will be used in case there is an opportunity to combine the previous time patterns with the likelihood of a high or low occurring.
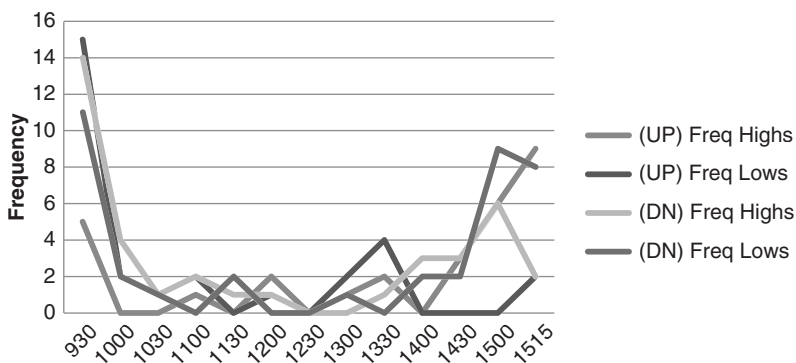
S&P 500 Highs and Lows

The pattern shown for the S&P futures in Figure 15.7 is typical of most markets. The open is the time of day with the greatest chance of being the day's high or low price, twice as likely as the closing time and eight times more likely than near the middle of the trading session. This pattern should not be a surprise because it has not changed in 50 years. What could be a surprise is that the close of the day is less likely to be a high or low than the open. We often remember strong or weak closes, rather than strong or weak opens.

Can this be turned into a trading opportunity? Once the market has been open for an hour and trades away from the opening price, that early high or low becomes a strong candidate for a daily high or low. Then, if that level is broken, the close becomes the next likely time for a high or low. The concept of a new high or low after midday could be the basis for an intraday breakout signal.

This pattern could change if there is a strong or weak gap open. Figure 15.8 shows the frequency of highs and lows following a gap of ±2.0%. There are two significant differences. First, the difference in the frequency of the open and close is much smaller. Instead of 2.5:1, as in the first chart, it is now about 1.6:1. More important, the other times of day are much less likely to post a high or low. For the case of the strong upwards gap, the low of the day is overwhelmingly going to be at the open. Consistent with that is the high of the day, which is more likely to be at the close than the open.

EURUSD Highs and Lows

Because currencies are actively traded around the world, the pattern for EURUSD is different from the S&P. In the previous section, the analysis of time-of-day patterns showed that activity could be separated into three parts: from the open of Europe to the open of

FIGURE 15.8 S&P intraday highs and lows when there is a gap open, higher or lower, exceeding 2.0%.

the U.S., from the open of the U.S. to the close of Europe, and from the close of Europe to the close of the U.S. The same is true for the frequency of highs and lows. From 2 A.M. to 8 A.M. the frequency of EURUSD highs and lows is very similar to the S&P chart. This is when Europe is the dominant player. The difference between this and the S&P chart is that the early open is far more likely to be the high or low of the day. From 8 A.M. to 10 A.M. both Europe and the United States are open and the frequency of both highs and lows are essentially sideways. Once Europe closes, the normal pattern repeats, but in a more subdued manner (see Figure 15.9).

Using the Intraday High and Low Patterns for Trading

There are few simple rules that come out of the intraday high-low patterns that could help trading. Because there is a high frequency of the high or low occurring during the first bar
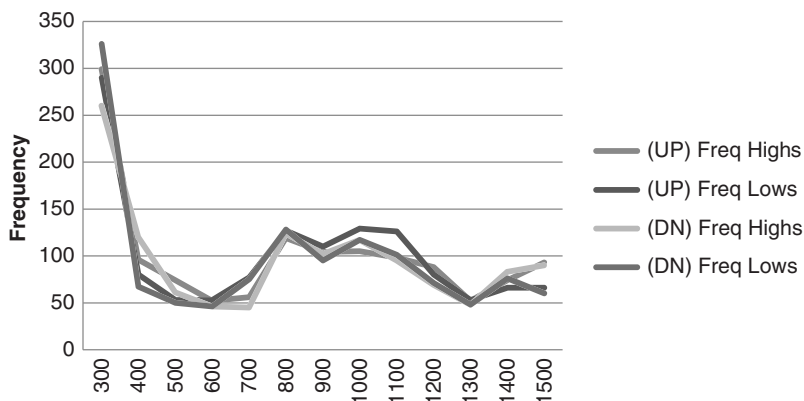
FIGURE 15.9 Frequency of intraday highs and lows for EURUSD futures, 2000 through May 2011.

of the day, you become more confident as the day progresses that an early extreme is a high or low. For example, the S&P gaps up 12.00 points after a bullish employment report. It then runs up to a high of up 15.00 points at the end of the first 30 minutes, then falls back to trade up 11.00 points by the end of the second bar. Have we seen a high or a low? It may still be too early to tell. If prices continue lower and make new lows, then we are much more confident that the high was made during the first bar. If prices make a new high, then we are equally confident that we have seen the low of the day. As each bar passes, we are more confident that we have seen the high or low. We can then trade in the direction of the high or low not yet seen, expecting that it will occur on the close. The sooner we make this decision, the more profit potential we can capture. A breakout of a high or low later in the day increases the chance of a close in that direction. Always place a stop-loss at the point where a new low would occur if you are long, or a new high if you have sold short.

Midday offers another trading opportunity because of the small chance of seeing a new high or low. If you have decided that the low occurred near the open, then buying on a test of the lows during the quiet midday session could provide good entry timing.

In an attempt to keep similar topics together, this strategy will be reviewed in more detail in the next chapter on intraday trading. At that point, the S&P and EURUSD markets will be tested with a strategy based on the distribution of highs and lows.

An exception to every rule

During the 1990s, the stock market experienced one of the greatest bull markets in history. From 1995 through 1996 nearly everyone was talking about the 20% annual returns that were easily achieved in the stock market. This affected the intraday patterns. Investors would most often buy on the open of trading, but they would also buy on any pullback during the day. The result was that the *U*-shaped pattern that we clearly see in Figures 15.7 through 15.9 were replaced by a jump at the open and a straight line throughout the rest of the day.

A Program for Intraday Highs and Lows

A program that prints a table of the frequency of intraday highs and lows is *TSM Intraday High-Low* and can be found on the Companion Website. It also allows for setting a gap threshold to filter out those days with small opening gaps which might be considered of uncertain direction.

OPENING GAPS

To any trader, an opening gap looks like an opportunity for a profit. Certainly, a large *gap opening* implies greater volatility for the day, and we expect a greater reaction. Based on the way the market closes, it may also forecast the direction of the next opening price. At least we would expect a volatile open to offer more information than a quiet open. Using only daily data, a gap analysis was performed to study the combination of patterns that followed a gap opening of various strength or weakness. A strong open followed by a stronger

close would allow a trader to buy on a pullback after the open. A consistently weak close after a strong open would be an opportunity to sell short near the open. Specific entry timing can be improved by using the patterns discussed in the previous sections.

The output for these tests are shown in Table 15.2 (a) and (b). Part (a) shows the frequency of occurrence of ABX (Barrick Gold) gaps from January 2000 through May 2011 based on increments of 25 basis points, shown on the left. In the center row are then the number of days that prices opened unchanged from the previous close. Below that are the days when prices gapped lower on the open, and above are the days prices gapped higher. The top and bottom rows have all those days when prices gapped greater than and less than 2.5% from the previous close. Part (b) uses the same values but expresses them as a percentage of the number of cases on each line. The percentages are most useful, but we also need to know the frequency to assess the significance. If there are only three cases on a line, the results are not very helpful.

Both parts (a) and (b) of Table 15.2 show the following columns:

1. The opening gap expressed in basis points (e.g., 150 is 1.50%).
2. The number of cases for that opening gap. The 150 line would include 150 up to 175 but not including 175.
3. Three columns explaining the closing price of the day:
   a. *Cont Dir.* The close continued (extended the move) in the direction of the gap.
   b. *Below Open.* The close was between the gap opening and the prior close.
   c. *Rev Dir.* The close reversed the direction of the opening gap, that is, if the open was higher, the close was lower.
4. The next three columns show the trading range pattern.
   a. *Cross Prev Close.* Prices crossed the prior closing price at least once during the trading session.
   b. *Adj from Open.* Prices reversed from the open but did not cross the prior close.
   c. *Cont Only.* Prices continued in the direction of the open and never reversed.
5. The last column shows the frequency of prices continuing on the next open in the direction of today's gap.

Gaps in Futures Markets

Although the example above used ABX, we will start with gaps in futures markets, the most popular being the *e*mini S&P contract. Instead of showing the detail in a comparable table, Figure 15.10 compares the frequency of gaps in ABX with those in the *e*mini S&P. While ABX has most of its gaps greater and less than 1%, as shown by the peaks at both ends, the S&P is exactly the opposite, with essentially no gaps greater or less than 1%. How can that happen? We know that the stock market has gapped higher and lower fairly often.

TABLE 15.2(a)   ABX Gaps in Increments of 25 bps

Gap	Cases	Closing Prices			Trading Range			Cont Next Day
		Cont Dir	**Below Open**	**Rev Dir**	**Cross Prev Close**	**Adj from Open**	**Cont Only**	**Cont Next Day**
250	131	62	44	25	40	85	5	68
225	48	27	18	3	6	39	2	27
200	60	31	15	13	24	34	1	31
175	85	39	25	19	36	47	2	41
150	131	66	33	29	59	59	9	75
125	151	77	32	40	71	66	9	95
100	167	85	26	52	93	62	7	92
75	222	103	34	75	146	57	16	107
50	240	102	17	112	184	25	25	145
25	246	114	7	115	218	10	14	146
0	196	0	0	0	0	0	0	0
−25	191	102	6	78	166	10	13	79
−50	228	117	13	83	176	29	11	97
−75	159	79	24	54	100	36	16	71
−100	157	74	38	42	93	50	6	70
−125	97	61	16	20	45	42	6	35
−150	89	48	18	19	40	43	4	42
−175	49	26	9	14	22	23	3	23
−200	47	29	11	7	13	31	3	18
−225	34	15	13	6	17	16	1	18
−250	94	57	23	11	20	68	5	37

TABLE 15.2(b)   ABX Results in Percentages

| Gap	Cases	Closing Prices			Trading Range			Cont Next Day
		Cont Dir	**Below Open**	**Rev Dir**	**Cross Prev Close**	**Adj from Open**	**Cont Only**	**Cont Next Day**
250	131	47.33	33.59	19.08	30.53	64.89	3.82	51.91
225	48	56.25	37.50	6.25	12.50	81.25	4.17	56.25
200	60	51.67	25.00	21.67	40.00	56.67	1.67	51.67
175	85	45.88	29.41	22.35	42.35	55.29	2.35	48.24
150	131	50.38	25.19	22.14	45.04	45.04	6.87	57.25
125	151	50.99	21.19	26.49	47.02	43.71	5.96	62.91
100	167	50.90	15.57	31.14	55.69	37.13	4.19	55.09
75	222	46.40	15.32	33.78	65.77	25.68	7.21	48.20
50	240	42.50	7.08	46.67	76.67	10.42	10.42	60.42
25	246	46.34	2.85	46.75	88.62	4.07	5.69	59.35
0	196	0	0	0	0	0	0	0
−25	191	53.4	3.1	40.8	86.9	5.2	6.8	41.4
−50	228	51.3	5.7	36.4	77.2	12.7	4.8	42.5
−75	159	49.7	15.1	34.0	62.9	22.6	10.1	44.7
−100	157	47.1	24.2	26.8	59.2	31.9	3.8	44.6
−125	97	62.9	16.5	20.6	46.4	43.3	6.2	36.1
−150	89	53.9	20.2	21.4	44.9	48.3	4.5	47.2
−175	49	53.1	18.4	28.6	44.9	46.9	6.1	46.9
−200	47	61.7	23.4	14.9	27.7	66.0	6.4	38.3
−225	34	44.1	38.2	17.7	50.0	47.1	2.9	52.9
−250	94	60.6	24.5	11.7	21.3	72.3	5.3	39.4
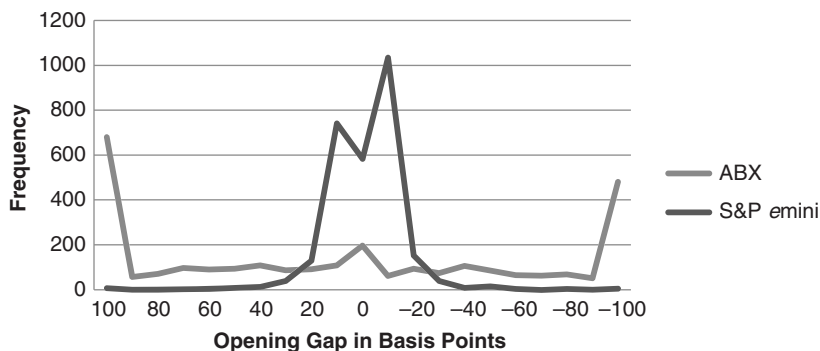
FIGURE 15.10   Comparison of gap frequency for ABX and the *e*mini S&P.

The answer is that the *e*mini S&P trades on Globex from 6 P.M. to the following 4:15 P.M. (New York time). Because the new session opens only 1 hour, 45 minutes after the previous session closes, and there is rarely any news to surprise the market, most gaps will occur on Sunday evening when the new week begins. Otherwise, the new opening price is very close to the previous closing price. To get a pattern similar to ABX we would need to look at the floor session for the big S&P contract, which has been declining in volume. Alternatively, we could set the time for the *e*mini session to only look at prices from 9:30 A.M. when the stock market opens. In the next section, we will look at gaps for individual stocks, which will have great similarity to the index.

A program, *TSM Gap Study*, that produces the gap tables can be found on the Companion Website.

Five futures markets, U.S. 30-year bonds, the German DAX, the EURUSD currency, gold, and wheat were studied, and a comparison of the frequency of gaps is shown in Figure 15.11. The data was from January 2000 through May 2011, continuous futures,
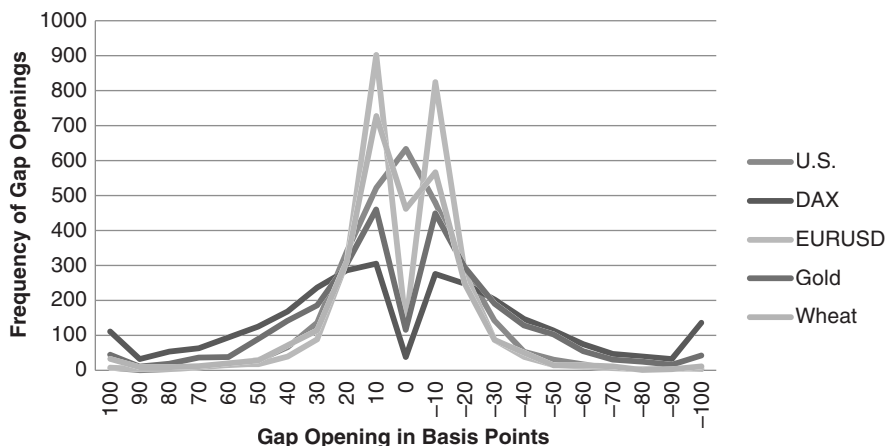
FIGURE 15.11   Frequency of gap opening in basis points for five futures markets, from January 2000 through May 2011.

which makes some of the percentage results slightly off due to back-adjusting. In general, they all follow a normal distribution, with the tails curving up at the end when gaps exceeded +1% and –1%. Of the five markets, the DAX is much more volatile than the others, with fewer cases in the middle and more in the tail. The EURUSD is at the other extreme, with more in the middle and fewer in the tail. U.S. bonds (US) is the only market that shows the unchanged cases greater than the small gaps on either side of the center and very little in the tail. This can be explained because the futures markets open at 8:20 A.M. (New York), 10 minutes ahead of the economic reports. On reporting days the market often opens very quiet in anticipation of the report. Had we moved the opening time to 8:40 A.M. using intraday data then there would be a large number of gaps in the tails.

Tables 15.3 (a) through (e) show the statistics for the five futures markets, each expressed as a percentage of the number of cases in the row and all from January 2000 through May 2011.

Looking at the same statistics across the five markets is a helpful way of understanding the patterns and deciding if there is an opportunity for a trading profit. Figure 15.12 shows the results of prices continuing the next day, on the open, in the same direction as today's gap. Only two markets stand out as different from the others, U.S. bonds and the EURUSD. Bonds have a high probability of continuing when the gap is +70 bp but nearly the opposite when the gap is +80 bp. However, the conclusion is based on too little

TABLE 15.3(a) U.S. 30-Year Bonds

		Closing Prices vs. Open			Trading Range			
Gap	Cases	Cont Gap Dir	Below Open	Rev Gap Dir	Crossed Prev Close	Pullback from Open	No Pullback	Cont Next Day
100	7	71	29	0	43	57	0	57
90	0	0	0	0	0	0	0	0
80	3	100	0	0	0	100	0	33
70	8	25	38	0	50	50	0	88
60	19	47	5	42	63	37	0	63
50	29	45	24	24	55	28	10	62
40	66	56	9	26	61	30	2	62
30	136	45	13	34	74	19	1	51
20	340	49	5	36	82	9	3	53
10	522	53	0	39	87	2	5	50
0	633	0	0	0	0	0	0	0
−10	479	48	0	46	87	1	4	48
−20	289	42	6	46	86	8	2	51
−30	142	50	9	39	74	20	1	45
−40	53	51	21	21	58	30	6	47
−50	30	50	20	20	53	37	0	40
−60	17	24	18	59	82	12	6	53
−70	6	83	0	17	33	67	0	50
−80	3	33	67	0	33	33	0	0
−90	5	40	20	40	80	20	0	20
−100	4	75	25	0	0	100	0	25

TABLE 15.3(b) German DAX

Gap	Cases	Closing Prices vs. Open			Trading Range			Cont Next Day
		Cont Gap Dir	Below Open	Rev Gap Dir	Crossed Prev Close	Pullback from Open	No Pullback	
100	111	64	21	14	35	60	5	49
90	32	38	28	34	56	44	0	59
80	54	54	20	26	52	46	2	39
70	63	41	21	37	57	40	2	49
60	94	49	19	30	60	37	3	45
50	125	47	18	34	60	37	2	50
40	168	54	16	30	65	32	2	46
30	237	48	14	38	78	22	0	49
20	286	54	7	39	83	14	3	58
10	305	55	3	40	91	5	3	54
0	38	0	0	0	0	0	0	0
−10	276	50	3	46	94	3	3	46
−20	248	42	9	48	86	12	1	42
−30	203	48	12	38	81	16	1	45
−40	147	49	11	39	75	24	1	48
−50	114	53	18	28	65	33	2	42
−60	75	55	20	25	60	35	4	51
−70	47	55	15	30	60	40	0	43
−80	40	58	8	35	45	48	8	43
−90	32	50	34	16	34	59	6	38
−100	136	44	33	22	43	57	1	48

TABLE 15.3(c) EURUSD Currency

Gap	Cases	Closing Prices vs. Open			Trading Range			Cont Next Day
		Cont Gap Dir	Below Open	Rev Gap Dir	Crossed Prev Close	Pullback from Open	No Pullback	
100	7	29	71	0	14	71	14	57
90	4	25	50	25	100	0	0	50
80	3	67	0	33	67	33	0	0
70	8	38	25	25	63	38	0	63
60	15	53	20	27	47	53	0	53
50	17	41	18	41	82	18	0	53
40	39	51	23	26	67	31	0	33
30	88	48	16	36	69	22	5	65
20	310	48	9	43	84	14	1	50
10	902	53	4	42	94	3	1	54
0	130	0	0	0	0	0	0	0
−10	824	51	2	45	92	4	2	50
−20	277	45	8	45	85	14	1	44
−30	87	33	21	45	83	16	1	48
−40	38	50	13	37	82	18	0	55
−50	15	53	20	27	73	27	0	67
−60	10	70	20	10	40	60	0	10
−70	6	83	0	17	67	33	0	33
−80	2	50	0	50	50	50	0	0
−90	4	25	75	0	50	50	0	25
−100	5	20	60	20	60	40	0	40

TABLE 15.3(d) Gold

Gap	Cases	Closing Prices vs. Open			Trading Range			Cont Next Day
		Cont Gap Dir	Below Open	Rev Gap Dir	Crossed Prev Close	Pullback from Open	No Pullback	
100	45	49	24	24	38	60	2	42
90	11	45	27	27	36	55	9	64
80	19	42	16	37	47	47	5	63
70	37	65	22	14	43	49	8	57
60	38	58	21	18	53	42	5	55
50	90	58	21	19	43	50	7	52
40	142	58	24	18	50	40	8	47
30	186	49	15	34	70	26	4	48
20	304	56	8	32	72	20	6	54
10	460	52	3	41	87	5	5	50
0	116	0	0	0	0	0	0	0
−10	449	42	2	52	94	3	2	51
−20	295	41	11	46	80	15	2	52
−30	191	48	18	32	63	33	1	51
−40	129	50	25	24	51	47	0	49
−50	103	41	31	26	42	54	0	53
−60	55	49	35	15	42	55	0	49
−70	31	48	29	23	39	58	0	58
−80	25	40	40	16	28	72	0	48
−90	16	50	31	13	31	63	0	63
−100	43	42	44	12	28	72	0	51

TABLE 15.3(e) Wheat

Gap	Cases	Closing Prices vs. Open			Trading Range			Cont Next Day
		Cont Gap Dir	Below Open	Rev Gap Dir	Crossed Prev Close	Pullback from Open	No Pullback	
100	32	47	22	28	63	34	3	75
90	9	33	33	33	67	33	0	44
80	9	67	22	11	44	56	0	56
70	12	42	25	33	83	17	0	67
60	20	65	5	30	45	45	5	65
50	28	64	4	32	64	32	4	61
40	71	46	17	35	70	23	3	65
30	114	53	7	39	79	16	3	71
20	312	50	5	42	78	12	5	58
10	727	46	2	48	87	3	6	58
0	462	0	0	0	0	0	0	0
−10	566	56	2	38	89	3	4	44
−20	246	52	5	39	83	10	4	56
−30	87	51	3	41	72	21	5	40
−40	50	54	12	32	72	20	4	44
−50	14	57	7	36	57	29	0	64
−60	14	64	14	21	71	29	0	50
−70	11	45	0	55	73	18	9	36
−80	1	100	0	0	100	0	0	0
−90	3	67	33	0	0	100	0	33
−100	11	55	36	9	64	36	0	36
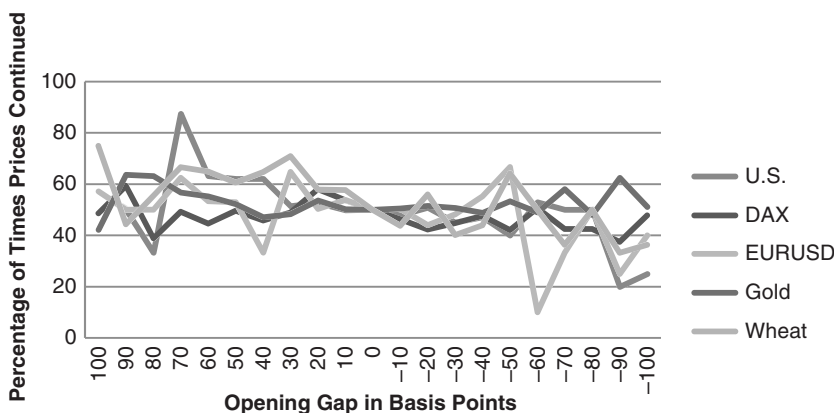
FIGURE 15.12 Likelihood of continuing in the same direction on the next open.

data, which makes it unreliable to trade. However, when there is a large downward gap in either bonds or euros, there is a low probability of continuing the next day, as seen in the lower right part of the chart. In addition, wheat and the DAX have the same pattern, although not as significant. Then any gap opening that shows a drop of at least 90 bp is likely to reverse on the next open. Unfortunately, there were only about 10 cases for each in 11 years; therefore, these results also lack confidence. If we look closely at wheat, we can see that all upwards gaps from 30 bps and higher had a 63% chance of continuing the next day, a more interesting outcome.

While the results of continuing the next day are not as clear as we would like, there are other values that have more potential for trading. For example, in looking at U.S. bonds, consider the two columns, *Closing below the open* and *Crossing the previous close*, and the rows that include gaps from +40 to –30 bps. That area of the table has been extracted below:

Gap	Cases	Closing Prices vs. Open	Trading Range
		Below Open	**Crossed Prev Close**
40	66	–9	–61
30	136	–13	–74
20	340	–5	–82
10	522	0	–87
0	633	0	0
–10	479	0	–87
–20	289	–6	–86
–30	142	–9	–74

For bonds, the chance of a small upwards gap closing below the open, or a downwards gap closing above the open, is very low. In addition, the chance of either an upward or downward gap crossing the previous close is high. We can then expect that an

upwards gap that retraces below the previous close should be bought with expectations of a close that will be above the open. The opposite would be true of a lower gap opening. This is really a confirmation that interest rate markets trend. Any unexpected move to the upside will close in that direction and the same for the downside. The results for the DAX are not at all the same, but the results for the EURUSD are very similar.

Do Individual Stocks Show Unique Gap Patterns?

Earlier there was a chart of the frequency of *e*mini S&P gaps compared to ABX, showing that ABX had a large number of gaps greater than 1%. Now we will look at more detail for ABX, XOM (Exxon-Mobile), AMZN (Amazon), and LUV (Southwest Airlines), shown in Tables 15.4(a) through (d). The data is all from January 2000 through May 2011 and the gaps are in increments of 25 basis points.

As with futures, it seems that the most interesting combinations are smaller gap openings that persist in closing in the direction of the gap after crossing through the previous closing price. There seems to be a low chance of a reversal during the trading day, but there are no numbers that indicate that the direction will continue the next day. We can try to understand that in terms of the analysis in Chapter 2, which concluded that the equity markets have the highest level of noise.

A Simple Gap Trading Method for Stocks

A basic approach to trading gaps was taken by Reverre.[5] Beginning with a 1-year study of the distribution of opening gaps for General Electric, Reverre noted that the distribution was very narrow, with 85% of the days opening with less than a 1% gap, and only 5% of the days showing gaps greater than 2%. Apply the following rules:

- Place limit orders to buy if prices open more than 1% lower, and sell if they open more than 1% higher, than the previous close.
- Exit the trade at the close of the day.

Although there may be additional profits from holding the trade for more than one day, this simple method produced small but consistent profits when applied to the Dow stocks. Reverre makes two important observations:

1. Trading the medium-sized gap of 1% showed consistency because gaps often represented noise and were likely to correct.
2. Trading gaps greater than 2% gave much larger profits and losses. Larger gaps were more likely to be structural and may not correct.

---

[5]Stéphane Reverre, "Trading the Opening Gap," *Technical Analysis of Stocks & Commodities* (November 1999); and Stéphane Reverre, "Trading against the Gap," *Technical Analysis of Stocks & Commodities* (Bonus Issue 2003).

| TABLE 15.4(a) | ABX (Barrick Gold) |

		Closing Prices			Trading Range			
Gap	Cases	Cont Dir	Below Open	Rev Dir	Cross Prev Close	Adj from Open	Cont Only	Cont Next Day
250	131	47	34	19	31	65	4	52
225	48	56	38	6	13	81	4	56
200	60	52	25	22	40	57	2	52
175	85	46	29	22	42	55	2	48
150	131	50	25	22	45	45	7	57
125	151	51	21	26	47	44	6	63
100	167	51	16	31	56	37	4	55
75	222	46	15	34	66	26	7	48
50	240	43	7	47	77	10	10	60
25	246	46	3	47	89	4	6	59
0	196	0	0	0	0	0	0	0
−25	191	53	3	41	87	5	7	41
−50	228	51	6	36	77	13	5	43
−75	159	50	15	34	63	23	10	45
−100	157	47	24	27	59	32	4	45
−125	97	63	16	21	46	43	6	36
−150	89	54	20	21	45	48	4	47
−175	49	53	18	29	45	47	6	47
−200	47	62	23	15	28	66	6	38
−225	34	44	38	18	50	47	3	53
−250	94	61	24	12	21	72	5	39

| TABLE 15.4(b) | XOM (Exxon-Mobil) |

		Closing Prices			Trading Range			
Gap	Cases	Cont Dir	Below Open	Rev Dir	Cross Prev Close	Adj from Open	Cont Only	Cont Next Day
250	19	47	47	5	16	84	0	58
225	9	67	33	0	22	78	0	44
200	9	67	22	11	11	89	0	44
175	27	63	15	22	33	63	0	37
150	35	60	23	17	51	40	9	49
125	72	44	25	28	42	57	1	42
100	151	54	24	17	42	52	3	54
75	263	47	22	31	59	37	2	55
50	408	52	11	35	68	25	5	50
25	386	54	4	40	84	11	4	58
0	202	0	0	0	0	0	0	0
−25	375	47	5	46	89	6	3	46
−50	312	46	11	40	77	19	3	47
−75	234	47	21	29	63	36	0	52
−100	107	43	27	29	53	45	2	41
−125	75	41	36	17	45	52	1	48
−150	45	53	22	24	36	62	2	42
−175	25	52	24	24	40	60	0	60
−200	20	30	35	30	65	30	0	50
−225	10	50	50	0	50	50	0	60
−250	37	62	32	5	27	70	3	41

TABLE 15.4(c) AMZN (Amazon)

Gap	Cases	Closing Prices			Trading Range			Cont Next Day
		Cont Dir	Below Open	Rev Dir	Cross Prev Close	Adj from Open	Cont Only	
250	153	55	24	21	37	61	1	45
225	30	30	20	50	67	30	3	43
200	33	58	21	18	42	55	3	33
175	66	53	15	29	55	42	3	39
150	70	47	20	33	74	24	0	66
125	87	49	15	36	59	39	0	44
100	144	53	13	33	69	27	2	48
75	187	47	9	43	79	19	1	42
50	259	46	7	46	86	12	2	47
25	297	49	4	44	91	4	2	46
0	50	0	0	0	0	0	0	0
−25	286	48	1	48	92	5	2	48
−50	302	43	11	46	86	10	1	55
−75	217	46	14	39	73	24	2	45
−100	147	46	16	37	67	31	0	51
−125	93	51	22	28	63	33	1	53
−150	83	51	20	28	61	35	2	51
−175	71	52	20	28	62	34	1	52
−200	38	42	24	29	61	39	0	50
−225	37	43	27	30	54	41	0	62
−250	173	49	25	24	40	55	2	49

TABLE 15.4(d) LUV (Southwest Airlines)

Gap	Cases	Closing Prices			Trading Range			Cont Next Day
		Cont Dir	Below Open	Rev Dir	Cross Prev Close	Adj from Open	Cont Only	
250	74	42	32	19	45	54	0	59
225	20	50	25	10	50	40	0	50
200	38	47	32	18	45	53	3	42
175	48	54	21	23	48	40	4	40
150	56	43	25	30	59	34	4	54
125	95	44	24	24	54	35	4	53
100	111	50	19	26	63	26	4	53
75	196	49	12	36	68	22	5	51
50	304	47	4	44	75	13	7	45
25	301	49	1	46	83	2	10	45
0	283	0	0	0	0	0	0	0
−25	301	48	1	46	88	3	5	47
−50	299	49	12	36	76	14	6	56
−75	262	48	15	32	71	18	5	47
−100	143	52	20	24	55	34	6	51
−125	92	50	23	24	51	37	3	54
−150	58	55	16	26	41	50	7	47
−175	34	47	29	24	44	53	3	50
−200	34	38	38	15	53	41	3	50
−225	23	35	35	26	57	43	0	61
−250	51	43	35	20	29	67	4	47
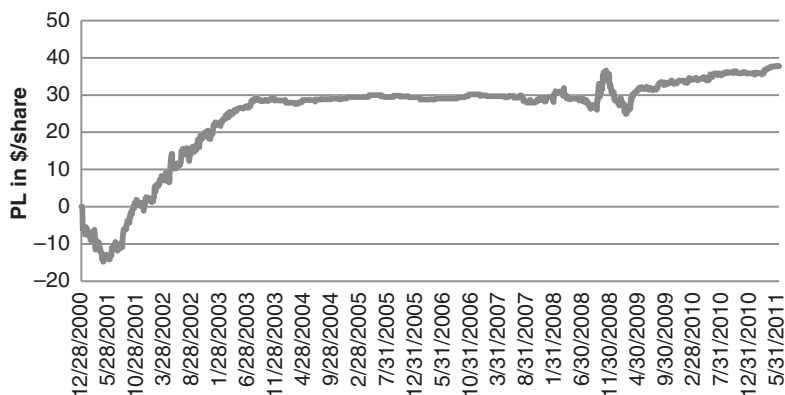
FIGURE 15.13   General Electric cumulative PL using a simple gap filter day trading method.

Test results showed that a 1% gap yielded an average long profit of $0.057 and average short profit of $0.034, which makes the brokerage commissions an important factor. When optimizing the gap size in 2004, 1.25% showed the best results; however, this is likely to change with the overall volatility of the market.

Updating this method, using GE data from January 2000 through May 2011 and a 2% gap filter, gave the cumulative profits, in $/share, shown in Figure 15.13. For a 1% filter, the program returned only 2.13¢ per share (without costs), and with a 2% filter it increased to 3.39¢ per share, with longs doing somewhat better than shorts. The period for mid-2003 until the subprime crisis in 2008 had very low volatility and few trades; however, if we remove that from the statistics, the profits per trade should increase significantly and make this simple approach a viable possibility. The program code is on the Companion Website as *TSM MR Gap System*.

Close-to-Close Gaps

Reverre also applied a similar approach to large 1-day moves, looking for a correction over the next one to three days. Using the largest 20% moves in General Electric, or about 1.28 standard deviations, during the period December 1995 to February 1999, a *buy* signal occurred when prices fell by more than 1.96% and a *sell* signal when prices rose by more than 2.24%. Although 70% of the corrections occurred by the close of the following day, best results came by holding the trade for two days. Entries could be placed as limit orders in advance of the close.

Tested over the 30 Dow stocks, maximum returns came from 1-day moves with gaps of 4% to 5%, holding for 1 day. A wider range of gaps, from 2.5% to 5% gave good results holding for 2 days. While less than half of the Dow stocks showed net profits, the overall profitability was good. Both long and short positions performed well, even though the test period was near the end of a strong bull market.

From 2000 to May 2011, the same tests were still profitable, but returned only 2 to 3¢ per share using the smaller gaps of about 2%. Larger gaps of 4% showed losses, which

means that limiting the gaps to between 2% and 4% could be highly profitable if you can decide whether a range of gaps is sensible or a case of overfitting. The program, *TSM Close-to-Close MR System* is available on the Companion Website.

WEEKDAY, WEEKEND, AND REVERSAL PATTERNS

There are many patterns that are important to traders because they seem intuitively correct. In Chapter 10, Seasonality and Calendar Patterns, we discussed month-end patterns, which are attributed to the flow of funds in and out of large pensions and investment firms as well as evening up and resetting of positions before an accounting period. The next three sections are also concerned with day-to-day patterns. The first, *weekday patterns*, looks only at the closing prices during the five days from Monday through Friday in order to find recurring close-only patterns. A week is often considered to have integrity as a single unit of time, and participants are accused of acting in the same way each week. During a bull market, we are made to believe that new investors will enter on Monday; by Tuesday or Wednesday the inflows have waned, and prices fall back. This study attempts to isolate repeated patterns. For example, if prices have moved in the same direction from Monday through Thursday, should we expect a correction due to profit-taking on Friday? Because patterns represent human behavior, the results will be insightful.

Weekend patterns are considered independently in the second study. The opening direction on Monday could restore the trend of the prior week, if those traders who liquidate on Friday intend to reset their positions. The weekend is also an extended period for unexpected news or a build-up of public interest. In a weather market, the fickleness of nature can cause traders to take defensive action ahead of the weekend. This study will also attempt to relate the direction of the Monday opening price to some pattern or trend of the preceding week.

Reversal patterns are not based on a day of the week but have often been discussed as a leading indicator. They may also be used as a filtering or timing device.

Weekday Patterns

During their constant exposure to the market, professional traders often observe patterns in weekly price movement; their acceptance of these patterns is as old as the market itself. In *Reminiscences of a Stock Operator*, the fictional character Larry Livingston (who we believe to be Jesse Livermore) begins his career recording prices on a chalk board above the floor of the New York Stock Exchange, eventually becoming aware of patterns in these sequences of numbers. The most accepted pattern is the *Tuesday reversal*, which is taken as commonplace by close observers of the market. When questioned why a strong soybean market on the first of the week is reversed the next day, an aged member of the Board of Trade would shrug his shoulders and quote: "Up on Monday, down on Tuesday." If this is true, there is a trading opportunity.

If a commonly accepted idea is not enough to be convincing, consider the additional idiosyncrasies of human behavior: The weekend allows a build-up of sentiment, which should result in greater activity on Monday. Coupled with adding back positions that were liquidated prior to the weekend, this may cause a disproportionate move on Monday, especially early in the session. This pattern may be further exaggerated when a clear trend exists. With this overbought or oversold condition on Monday, it is likely that Tuesday would show a correction. However, trying to justify how the market works is not as significant as looking at the numbers.

The first step is to define the weekday pattern. This was done in terms of the Friday-to-Monday move (close-to-close). Monday always received the value X, regardless of whether its direction from Friday was up or down. For each day that closed in the same direction as the Friday-to-Monday move, another X is used; when the close reversed direction, an O is recorded. Therefore, XOXXO means that Tuesday and Friday, represented by O, closed in the opposite direction from the prior Friday-to-Monday move, while Wednesday and Thursday were in the same direction. This could have meant either of the situations:

		(1)	(2)
Monday	X	Up	Down
Tuesday	O	Down	Up
Wednesday	X	Up	Down
Thursday	X	Up	Down
Friday	O	Down	Up

It might be that there is a distinction between the weeks that begin with an upward move on Monday rather than a decline, but both cases were combined. It would also be reasonable to have assumed that the upwards bias of interest rate markets during the past 10 years would justify a separation of patterns according to their initial direction. Additional tests might filter the results by overall market volatility or by larger-than-average volatility on Monday. Instead, this assumes that the pattern, rather than the direction, is most important and likely to be the robust solution. We will look at filtering after a basic view of weekday patterns.

A sampling of different stock and futures markets are including in the next sections. The stocks are Bank of America (BAC) and Intel (INTL), but also the *e*mini S&P futures as well as NASDAQ futures. Other futures markets are U.S. 30-year bonds, EURUSD currency, gold, heating oil, and wheat, a small sample of different sectors.

Results are shown as a binary tree with columns for each day of the week. Because all tests cover the period from January 2000 through May 2011, there are about 500 weeks. On Monday, all 500 are used, and on Tuesday the two possible directions, up and down, will split that number according to the frequency of occurrence. By Friday there are 16 possible outcomes, and, if they are all of equal probability, then each will have 31 entries. Because the number of entries drops quickly, more extreme probabilities earlier in the week would give greater confidence. These values were generated by the program *TSM Weekday Study*, available on the Companion Website.

*e*mini S&P

We start with the *e*mini S&P because it is one of the most liquid markets and represents the stock market as a whole. The first point to notice is that Tuesday has a 54% chance of reversing. While not as extreme as we would prefer, this confirms the belief "up on Monday, down on Tuesday." The period from 2000 through May 2011 have considerable market swings, but overall prices spent more time going up than down (even if the net result was lower), so this study may benefit from a trend filter.

In Figure 15.14 the most interesting pattern is on Friday, where a pattern of XOOO has a 57% chance of becoming XOOOX, where Friday moves in the direction of Monday. However, there is only a 6% chance this pattern will develop.

Bank of America and Intel

Because the S&P represents the overall market, we would expect some similarity in the pattern of the large cap stocks. Instead we find that BAC has very different patterns, with the most extreme on Friday following an XXOX pattern and favoring a reversal XXOXO by 66% as shown in Figure 15.15.

emini S&P, January 2000–May 2011

Monday	Tuesday	Wednesday	Thursday	Friday
			XOXX 51	XOXXX 49
				XOXXO 51
		XOX 52	XOXO 49	XOXOX 46
				XOXOO 54
	XO 54		XOOX 48	XOOXX 44
				XOOXO 56
		XOO 48	XOOO 52	XOOOX 57
				XOOOO 43
X 100			XXXX 47	XXXXX 54
				XXXXO 46
		XXX 49	XXXO 53	XXXOX 52
				XXXOO 48
	XX 46		XXOX 46	XXOXX 53
				XXOXO 47
		XXO 51	XXOO 54	XXOOX 52
				XXOOO 48
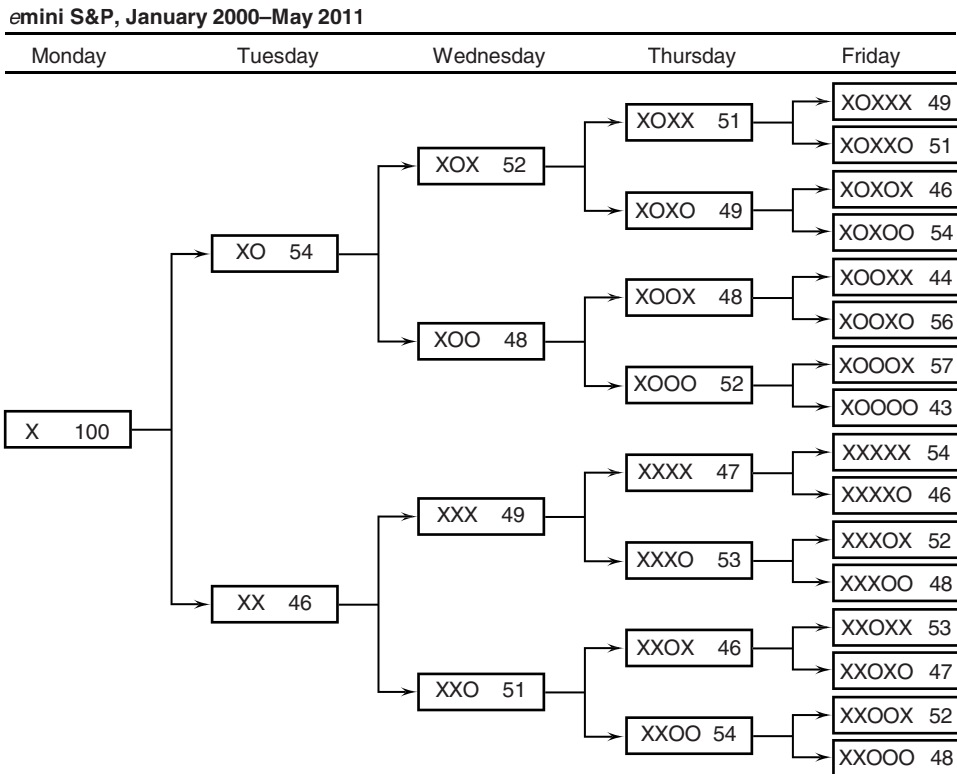
FIGURE 15.14 *e*mini S&P weekday patterns.

Bank of America (BAC), January 2000–May 2011

Monday	Tuesday	Wednesday	Thursday	Friday
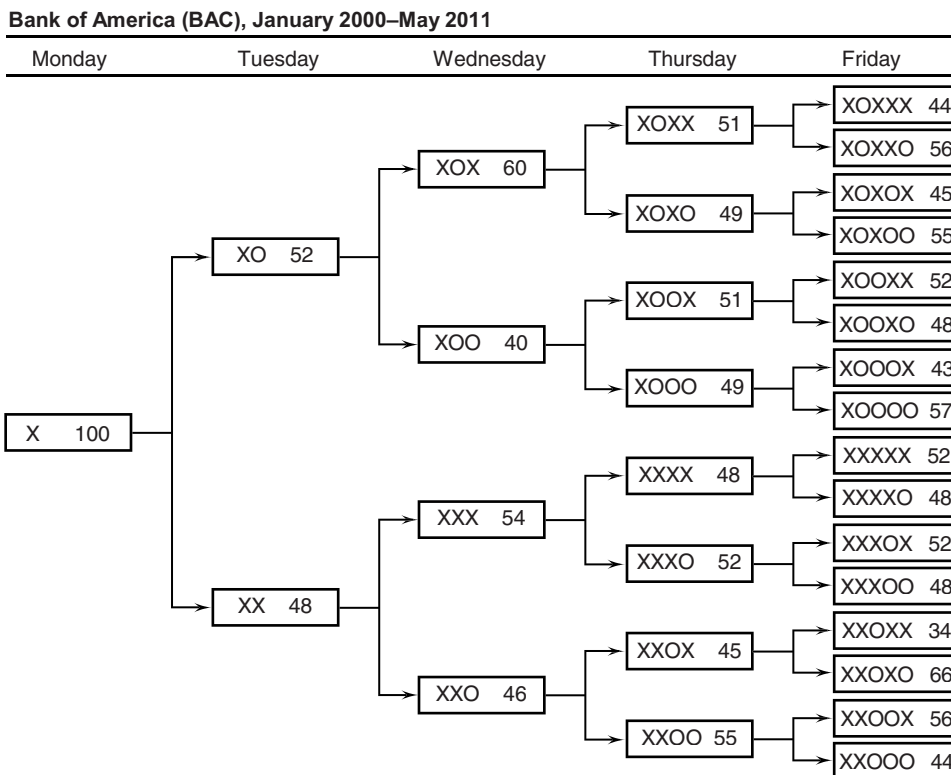
FIGURE 15.15   Bank of America weekday patterns.

Intel has a stronger tendency to reverse on Tuesday, but then follows the S&P pattern on Friday, with a 64% chance of producing an XOOOX pattern, shown in Figure 15.16.

Interest Rates and Commodities

The U.S. 30-year bond is representative of the long end of the yield curve and more volatile than shorter maturities. It has no interesting patterns through Wednesday, but on Thursday it has a good chance of a reversal if Tuesday and Wednesday went in the same direction as the Monday move, and if that fails, then Friday also has a good chance of a reversal (Figure 15.17).

The EURUSD is a very deep market, but has no strong patterns (Figure 15.18). The best seem to be on Thursday, where XXX is followed by another X 58% of the time, and XOX is followed by another reversal O about 57% of the time.

Gold has a number of patterns with a good bias beginning on Wednesday, when a continued move on Tuesday has a 55% chance of reversing on Wednesday (Figure 15.19). The best combination on Thursday favors flip-flopping, where a Tuesday reversal is followed by a Wednesday reversal which then has a 57% chance of reversing on Thursday, giving an up-down-up-down pattern. Friday has a number of strong cases, three of which reverse from Monday's move. That increases the chance of success on Friday.

Intel Corporation (INTL), January 2000–May 2011
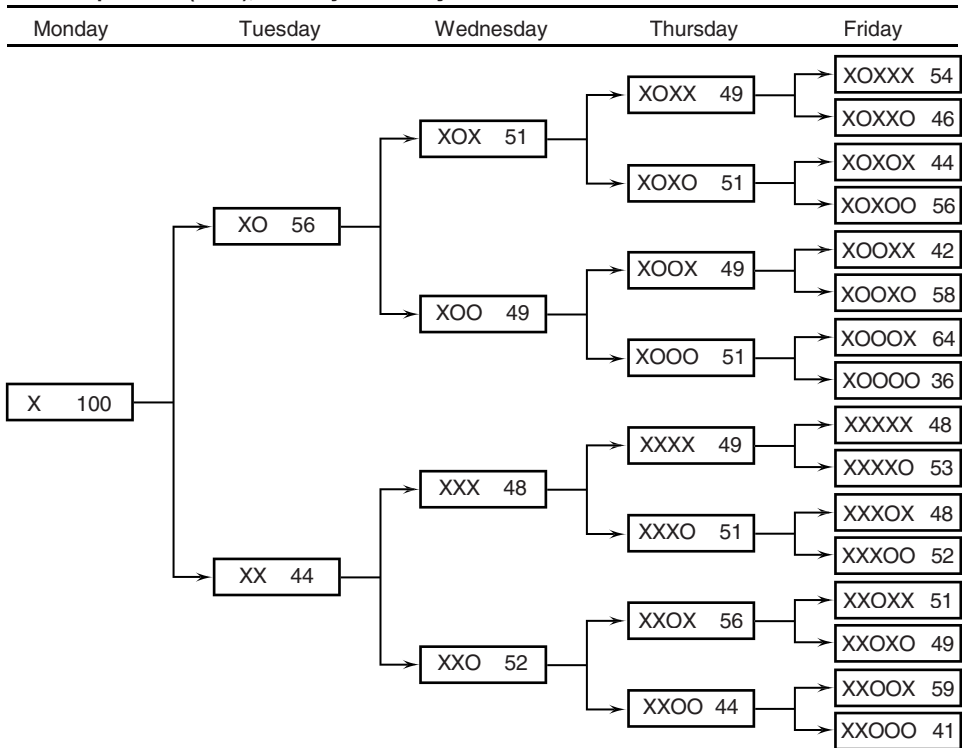
FIGURE 15.16 Intel weekday patterns.

U.S. 30–year bonds, January 2000–May 2011
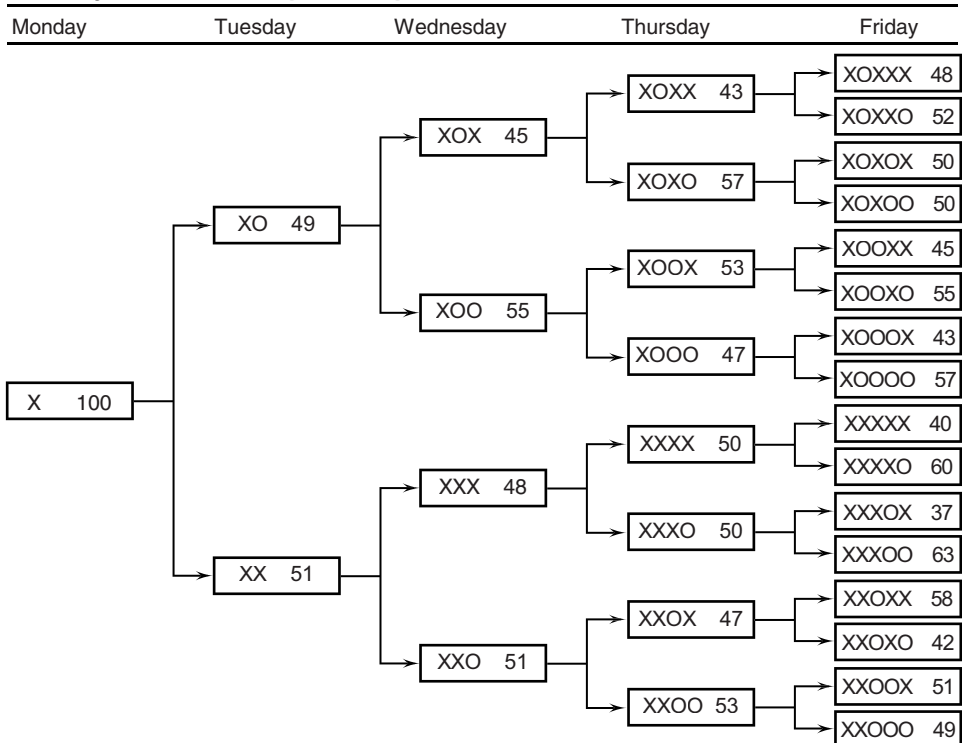
FIGURE 15.17 U.S. 30-year bonds weekday patterns.

Euro currency (EURUSD), January 2000–May 2011

Monday	Tuesday	Wednesday	Thursday	Friday
				XOXXX 55
			XOXX 43	XOXXO 45
		XOX 50		XOXOX 56
			XOXO 57	XOXOO 44
	XO 52			XOOXX 40
			XOOX 47	XOOXO 60
		XOO 50		XOOOX 56
			XOOO 53	XOOOO 44
X 100				XXXXX 56
			XXXX 58	XXXXO 44
		XXX 46		XXXOX 57
			XXXO 42	XXXOO 43
	XX 48			XXOXX 49
			XXOX 51	XXOXO 51
		XXO 54		XXOOX 44
			XXOO 49	XXOOO 56
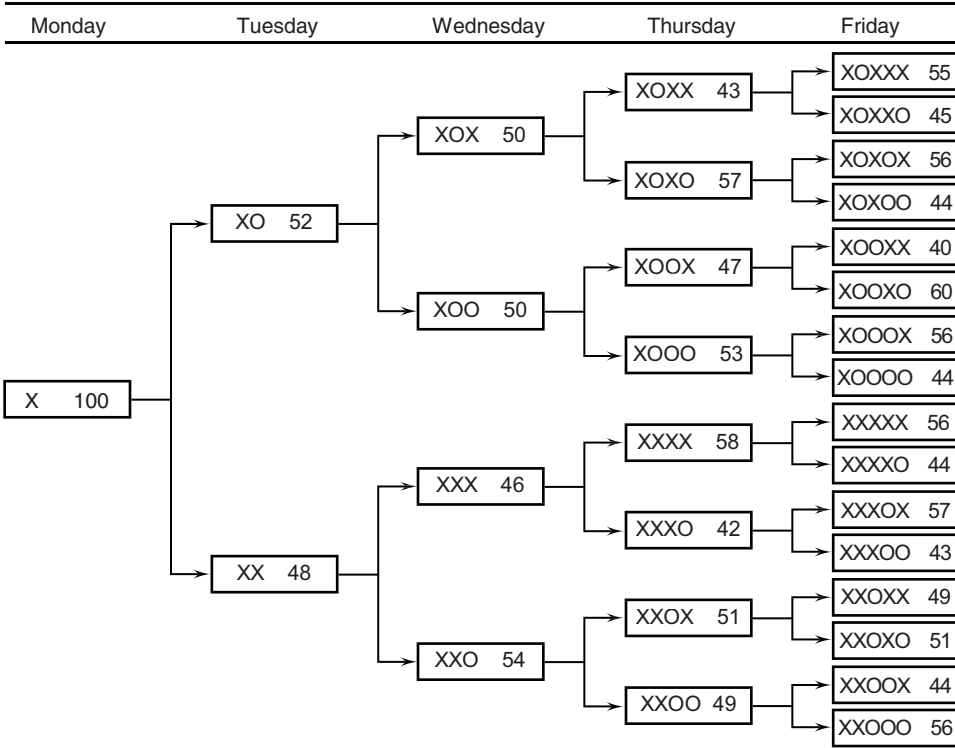
FIGURE 15.18 EURUSD weekday patterns.

Gold, January 2000–May 2011

Monday	Tuesday	Wednesday	Thursday	Friday
				XOXXX 50
			XOXX 43	XOXXO 50
		XOX 49		XOXOX 59
			XOXO 57	XOXOO 41
	XO 51			XOOXX 33
			XOOX 55	XOOXO 67
		XOO 51		XOOOX 59
			XOOO 45	XOOOO 41
X 100				XXXXX 58
			XXXX 52	XXXXO 42
		XXX 45		XXXOX 60
			XXXO 48	XXXOO 40
	XX 49			XXOXX 52
			XXOX 49	XXOXO 48
		XXO 55		XXOOX 45
			XXOO 51	XXOOO 55
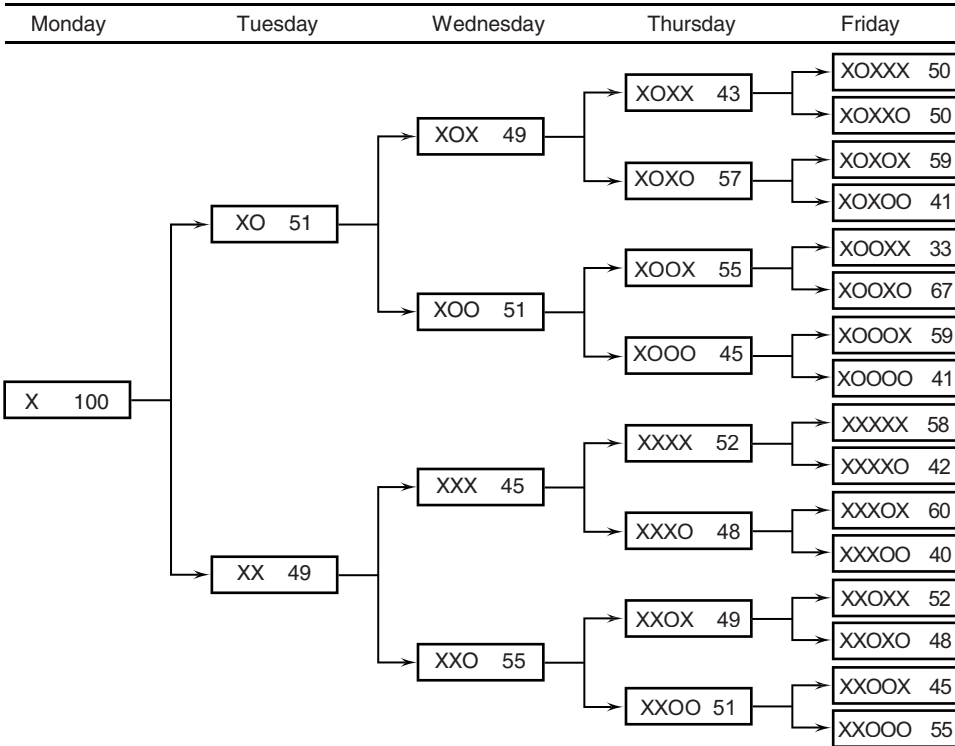
FIGURE 15.19 Gold weekday patterns.

Heating Oil, January 2000–May 2011
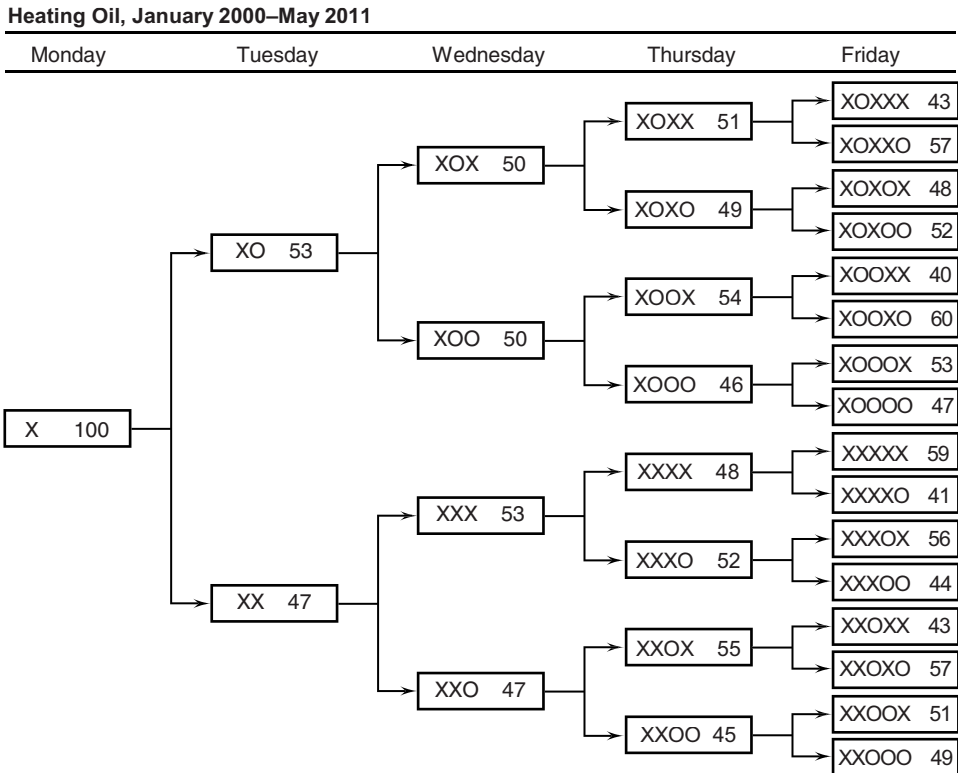
FIGURE 15.20 Heating oil weekday patterns.

Heating oil (Figure 15.20) is the most seasonal of the energy markets, but does not show any high-probability patterns until Friday. The patterns with the highest likelihood all reverse on Friday relative to Monday, except for the case where Monday through Thursday all went in the same direction; then Friday continued that direction.

Wheat is an interesting study. At first glance, the numbers are very normal, but collectively there are consistent patterns (Figure 15.21). Consider Thursday, where every pattern favors a reversal from the Monday direction. There is a similar case for Friday, but only when Tuesday shows a reversal. Then Friday favors a reversal for all patterns. When Tuesday does not show a reversal, then Friday favors the same direction as Monday.

Trading Weekday Patterns

The likelihood of a pattern repeating could always be improved by filtering the data by either a trend or volatility. We would expect more continuation patterns during a trend and perhaps more reversals when there is high volatility. Given the number of cases that remain on Friday without filtering, about 30 over the past 10 years, the filtered cases would provide a very small sample.

Instead, consider two cases:

1. Wheat, where every successful pattern for Thursday was a reversal of the Monday direction, or Friday where it reverses only when Tuesday shows a reversal.

Wheat, January 2000–May 2011
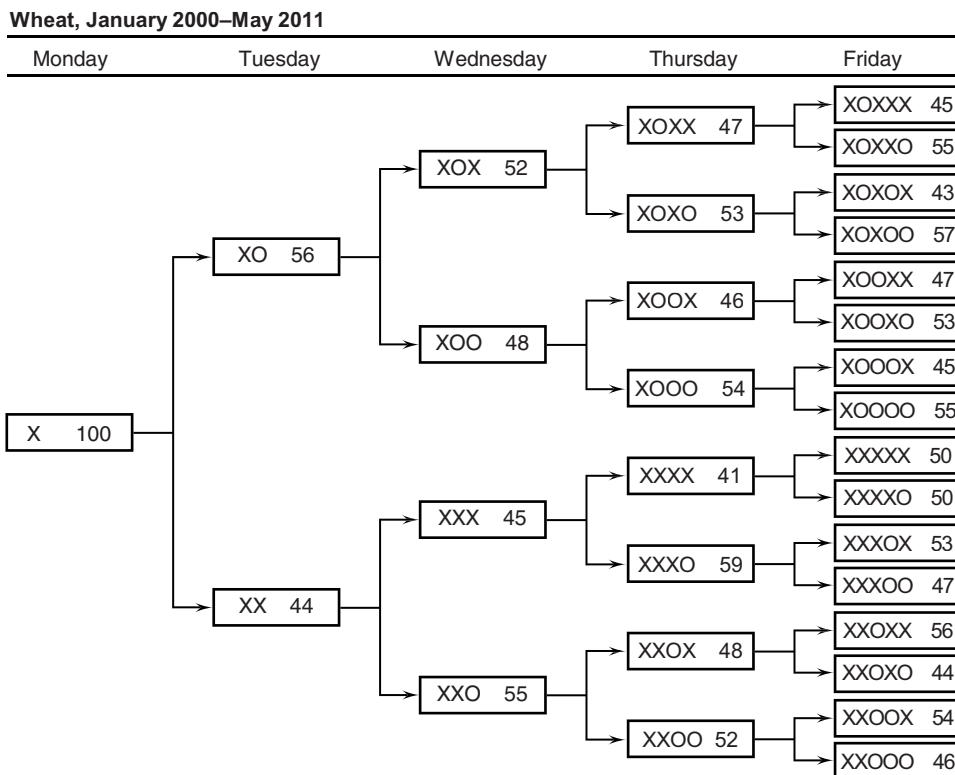
FIGURE 15.21 Wheat weekday patterns.

2. Bonds have a 55% chance of a XOO pattern. If Monday is up and Tuesday is down, we sell on Tuesday's close, expecting another down day. But Wednesday is up, giving us an XOX pattern instead. Thursday has a 57% chance of XOXO, then we can stay short and try again.

Your trading success is improved if you can find situations where one losing day has another good chance of turning into a profit simply by holding the same trade. This study could be taken further by looking at whether the open of the next day is a better point to enter a reversal trade than the previous close. Therefore, if Monday and Tuesday were both lower, a long entered on the open Wednesday may prove more profitable and less risky than a long entered on the close of Tuesday. A program to test weekday patterns is *TSM Weekday Trading*, on the Companion Website.

Weekday Patterns with a Trend Filter

Are the weekday patterns clearer when they are in context with the price trend? In a trending market, we would expect Monday to open in the direction of the trend, and most days during the week to move in the same direction; otherwise, the trend would not exist, or it would not be a profitable way to trade. But we know that trends are generally profitable, and longer calculation periods tend to be better than shorter ones. To verify that premise,

a trend filter was added to the study and the program called *TSM Weekday with Trend.* Three trend periods were tested, a 20-day, 50-day, and 200-day moving average. Patterns were recorded only if the move from Friday to Monday (close to close) was in the direction of the trendline (the trend today compared to the previous day's trend). As expected, the likelihood of certain patterns, and combinations of patterns, increased noticeably. Four markets are compared, the *e*mini S&P, U.S. 30-year bonds, gold, and EURUSD currency futures for the same 11-year period beginning January 2000. The tables that compare the original pattern with the trend-filtered patterns are shown in Tables 15.5 through 15.8.

TABLE 15.5   *e*mini S&P Weekday Patterns Filtered by Trends

	Number of Cases				**Percentage of Occurrence**			
Pattern	**No Trend**	**Trend 20**	**Trend 50**	**Trend 200**	**No Trend**	**Trend 20**	**Trend 50**	**Trend 200**
Monday	1	1	1	1	100%	100%	100%	100%
Tuesday								
XO	269	143	148	141	54	54	56	53
XX	226	121	118	127	46	46	44	47
Wednesday								
XOX	139	69	78	80	52	48	53	57
XOO	130	74	70	61	48	52	47	43
XXX	110	59	54	66	49	49	46	52
XXO	116	62	64	61	51	51	54	48
Thursday								
XOXX	71	34	40	43	51	49	51	54
XOXO	68	35	38	37	49	51	49	46
XOOX	63	42	37	29	48	57	53	48
XOOO	67	32	33	32	52	43	47	52
XXXX	52	31	32	32	47	53	59	48
XXXO	58	28	22	34	53	47	41	52
XXOX	53	28	28	24	46	45	44	39
XXOO	63	34	36	37	54	55	56	61
Friday								
XOXXX	35	17	25	23	49	50	63	53
XOXXO	36	17	15	20	51	50	38	47
XOXOX	31	13	12	13	46	37	32	35
XOXOO	37	22	26	24	54	63	68	65
XOOXX	28	20	17	16	44	48	46	55
XOOXO	35	22	20	13	56	52	54	45
XOOOX	38	18	19	18	57	56	58	56
XOOOO	29	14	14	14	43	44	42	44
XXXXX	28	21	22	17	54	68	69	53
XXXXO	24	10	10	15	46	32	31	47
XXXOX	30	13	10	18	52	46	45	53
XXXOO	28	15	12	16	48	54	55	47
XXOXX	28	13	15	15	53	46	54	63
XXOXO	25	15	13	9	47	54	46	38
XXOOX	33	21	22	19	52	62	61	51
XXOOO	30	13	14	18	48	38	39	49

TABLE 15.6 U.S. 30-Year Bonds Weekday Patterns Filtered by Trends

Pattern	Number of Cases				Percentage of Occurrence			
	No Trend	Trend 20	Trend 50	Trend 200	No Trend	Trend 20	Trend 50	Trend 200
Monday	1	1	1	1	100%	100%	100%	100%
Tuesday								
XO	242	126	134	130	49	48	50	48
XX	252	138	136	141	51	52	50	52
Wednesday								
XOX	108	61	59	56	45	48	44	43
XOO	134	65	75	74	55	52	56	57
XXX	124	66	69	76	49	48	51	54
XXO	128	72	67	65	51	52	49	46
Thursday								
XOXX	46	25	21	26	43	41	36	46
XOXO	62	36	38	30	57	59	64	54
XOOX	71	32	35	38	53	49	47	51
XOOO	63	33	40	36	47	51	53	49
XXXX	62	29	34	39	50	44	49	51
XXXO	62	37	35	37	50	56	51	49
XXOX	60	34	33	29	47	47	49	45
XXOO	68	38	34	36	53	53	51	55
Friday								
XOXXX	22	12	10	11	48	48	48	42
XOXXO	24	13	11	15	52	52	52	58
XOXOX	31	17	19	14	50	47	50	47
XOXOO	31	19	19	16	50	53	50	53
XOOXX	32	17	16	16	45	53	46	42
XOOXO	39	15	19	22	55	47	54	58
XOOOX	27	18	19	18	43	55	48	50
XOOOO	36	15	21	18	57	45	53	50
XXXXX	25	12	11	13	40	41	32	33
XXXXO	37	17	23	26	60	59	68	67
XXXOX	23	14	15	15	37	38	43	41
XXXOO	39	23	20	22	63	62	57	59
XXOXX	35	18	18	14	58	53	55	48
XXOXO	25	16	15	15	42	47	45	52
XXOOX	35	18	16	16	51	47	47	44
XXOOO	33	20	18	20	49	53	53	56

While there are individual patterns as high as 70% that stand out, combined results include many more cases and more opportunity. For the EURUSD, shown in Table 15.9, the 200-day trend filter increases the case where *Tuesday reverses, then Thursday reverses* from 55% to 61%. More important, the probabilities increase for all cases where the trend is used. Similarly, *if Tuesday does not reverse then Thursday does not reverse* is also increased everywhere, as is *if Tuesday reverses, then Friday does not reverse*. If we put this in context of what we understand about the trend, a Tuesday reversal is normal, but does not indicate great strength in the trend, so another reversal on Thursday seems

TABLE 15.7 Gold Weekday Patterns Filtered by Trends

Pattern	Number of Cases				Percentage of Occurrence			
	No Trend	Trend 20	Trend 50	Trend 200	No Trend	Trend 20	Trend 50	Trend 200
Monday	1	1	1	1	100%	100%	100%	100%
Tuesday								
XO	255	124	132	138	52	49	52	50
XX	239	129	122	138	48	51	48	50
Wednesday								
XOX	127	61	63	69	50	49	48	50
XOO	128	63	69	69	50	51	52	50
XXX	111	63	51	64	46	49	42	46
XXO	128	66	71	74	54	51	58	54
Thursday								
XOXX	55	24	26	23	43	39	41	33
XOXO	72	37	37	46	57	61	59	67
XOOX	60	29	30	31	47	46	43	45
XOOO	68	34	39	38	53	54	57	55
XXXX	64	40	33	40	58	63	65	63
XXXO	47	23	18	24	42	37	35	38
XXOX	65	34	40	37	51	52	56	50
XXOO	63	32	31	37	49	48	44	50
Friday								
XOXXX	30	13	16	15	55	54	62	65
XOXXO	25	11	10	8	45	46	38	35
XOXOX	40	19	22	28	56	51	59	61
XOXOO	32	18	15	18	44	49	41	39
XOOXX	24	12	17	10	40	41	57	32
XOOXO	36	17	13	21	60	59	43	68
XOOOX	38	21	24	24	56	62	62	63
XOOOO	30	13	15	14	44	38	38	37
XXXXX	36	19	20	21	56	48	61	53
XXXXO	28	21	13	19	44	53	39	48
XXXOX	27	11	9	11	57	48	50	46
XXXOO	20	12	9	13	43	52	50	54
XXOXX	32	18	22	20	49	53	55	54
XXOXO	33	16	18	17	51	47	45	46
XXOOX	28	12	15	16	44	38	48	43
XXOOO	35	20	16	21	56	63	52	57

reasonable. On the other hand, Friday goes back in the direction of the trend but is more pronounced if Tuesday reverses.

The S&P, the market followed by most traders, is not as clear as the EURUSD, and the results of combinations are not best in all cases based on the 50-day trend (Table 15.10). The first case, *Tuesday reverses*, confirms our intuition as well as market lore. In general, *if Tuesday does not reverse, then Thursday reverses*, but it is more likely that *if Tuesday reverses, then Friday reverses*. Again, finding combinations that are consistently better for all trends is a more robust solution.

TABLE 15.8   Euro Weekday Patterns Filtered by Trends

Pattern	No Trend	Trend 20	Trend 50	Trend 200	No Trend	Trend 20	Trend 50	Trend 200
Monday	1	1	1	1	100%	100%	100%	100%
Tuesday								
XO	250	133	128	127	51	50	50	48
XX	244	135	128	137	49	50	50	52
Wednesday								
XOX	122	65	62	63	49	49	48	50
XOO	128	68	66	64	51	51	52	50
XXX	110	65	63	66	45	48	49	48
XXO	134	70	65	71	55	52	51	52
Thursday								
XOXX	52	29	28	29	43	45	45	46
XOXO	70	36	34	34	57	55	55	54
XOOX	70	41	39	31	55	60	59	48
XOOO	58	27	27	33	45	40	41	52
XXXX	57	34	33	36	52	52	52	55
XXXO	53	31	30	30	48	48	48	45
XXOX	65	33	31	34	49	47	48	48
XXOO	69	37	34	37	51	53	52	52
Friday								
XOXXX	26	10	13	12	50	34	46	41
XOXXO	26	19	15	17	50	66	54	59
XOXOX	41	21	21	21	59	58	62	62
XOXOO	29	15	13	13	41	42	38	38
XOOXX	23	14	15	9	33	34	38	29
XOOXO	47	27	24	22	67	66	62	71
XOOOX	34	16	12	22	59	59	44	67
XOOOO	24	11	15	11	41	41	56	33
XXXXX	33	19	16	21	58	56	48	58
XXXXO	24	15	17	15	42	44	52	42
XXXOX	32	18	20	22	60	58	67	73
XXXOO	21	13	10	8	40	42	33	27
XXOXX	34	16	18	18	52	48	58	53
XXOXO	31	17	13	16	48	52	42	47
XXOOX	31	17	16	17	45	46	47	46
XXOOO	38	20	18	20	55	54	53	54

TABLE 15.9   EURUSD High-Probability Combinations Using Trends

	Trend Period			
	No Trend	20	50	200
If Tuesday reverses, then Thursday reverses	55.00	57.50	58.00	61.00
If Tuesday does not reverse, then Thursday does not reverse	54.50	57.50	60.50	56.50
If Tuesday reverses, then Friday does not reverse	51.75	52.00	60.00	55.25

TABLE 15.10 *e*mini S&P High-Probability Combinations Using Trends

	Trend Period			
	No Trend	20	50	200
Tuesday reverses	54.00	54.00	56.00	53.00
If Tuesday does not reverse, then Thursday reverses	53.50	51.00	48.50	56.50
If Tuesday reverses, then Friday reverses	52.75	55.50	57.25	55.00

Weekend Patterns

The weekend pattern relates how Monday opens and closes relative to the price moves of the previous week. In recording these patterns, three factors are considered:

1. *Friday liquidation.* Many short-term traders do not want to hold a position over the weekend, especially a 3-day weekend, due to the accumulation of unexpected news.

2. *Resetting positions on Monday.* Positions closed out on Friday may be reset Monday, especially during a trending market, a weather market, or when there is ongoing geo-political news.

3. *New positions entering Monday.* Longer-term investors use the weekend to digest the previous week's news and select those stocks or commodities to buy.

In the first two cases, there can be a large number of traders not willing to take weekend risk, even if they know that the Monday open is likely to add to their profits. They are trying to remove the less frequent, but large losses from their performance. This uncertainty can occur during agricultural weather markets, where a shortfall of rain has driven prices higher, but rain over the weekend would cause a sharp reversal. Recent geopolitical risk in the Middle East can also result in shifts of power over the weekend, a nasty and unpredictable situation for traders.

If there are enough investors and traders on the same side of the market, the 2-day weekend delay should result in large Monday price jumps and good profits for those willing to hold the trade over the weekend. Expecting that the best results will be found by combining the weekend patterns with a trend, or a high frequency weekday pattern on Friday, a number of different conditions were tested. These were limited to closing prices.

Trend Indicators
Prior Thursday-to-Friday price direction.
Friday-to-Friday (1-week) price direction.
Most frequent price direction last week (daily changes).

Based on each of these three trends and other patterns, the prices on Monday were tabulated to determine the frequency and consistency of the opening price, closing price, and whether the high or low prices on Monday were able to reach extremes that confirmed the same direction as each of the trends or patterns. Table 15.11 summarizes

TABLE 15.11 Weekend Results Conditioned on the Previous Week's Patterns

Pattern	emini S&P		U.S. 30-Year Bonds		Wheat		German DAX		Gold		Crude Oil		EURUSD	
	%Open	%Close	%Open	%Close	%Open	%Close	%Open	%Close	%Open	%Close	%Open	%Close	%Open	%Close
Thursday to Friday	41	45	40	43	48	47	51	50	49	47	47	52	53	45
Friday to Friday	37	47	37	46	45	48	51	53	47	49	48	49	52	44
Most frequent direction	38	49	40	45	45	47	55	53	51	47	48	51	51	42

TABLE 15.12 Monthly Price Moves Conditioned on Daily Patterns during the Previous Week

Pattern	emini S&P		U.S. 30-Year Bonds		Wheat		German DAX		Gold		Crude Oil		EURUSD	
	Cases	%Cont in Mon Dir	Cases	%Cont in Mon Dir	Cases	%Cont in Mon Dir	Cases	%Cont in Mon Dir	Cases	%Cont in Mon Dir	Cases	%Cont in Mon Dir	Cases	%Cont in Mon Dir
XOXXX	20	43	26	50	32	58	24	56	24	58	36	47	30	37
XOXXO	23	61	34	57	35	56	25	51	25	48	38	47	23	35
XOXOX	27	53	35	59	45	49	43	44	43	51	32	50	38	50
XOXOO	24	41	41	75	44	61	29	52	29	69	25	56	29	69
XOOXX	30	54	25	33	31	56	23	55	23	48	24	58	24	46
XOOXO	38	53	28	55	26	71	41	46	41	54	40	43	32	63
XOOOX	22	51	29	55	29	45	30	31	30	57	28	43	34	59
XOOOO	31	70	35	65	28	60	27	43	27	59	27	63	32	50
XXXXX	22	35	20	45	30	50	27	63	27	44	22	59	30	33
XXXXO	27	64	18	56	20	50	22	60	22	45	23	39	26	54
XXXOX	22	29	31	50	38	55	27	47	27	67	39	51	20	40
XXXOO	30	50	25	50	29	32	19	45	19	47	18	50	16	50
XXOXX	28	50	31	47	21	42	33	57	33	39	29	62	27	41
XXOXO	19	54	22	58	33	41	28	55	28	54	35	40	29	38
XXOOX	31	52	32	61	34	53	31	44	31	52	21	57	26	42
XXOOO	25	43	25	56	30	48	36	47	36	61	32	47	30	53
Same	3		6		6		4		6		6		3	
Reverse	5		2		3		4		2		4		7	

the results for seven markets, *e*mini S&P, U.S. 30-year bonds, wheat, German DAX, gold, crude oil, and the EURUSD forex futures. The period was January 2000 through June 2011. During this period, equity markets showed major trends and reversals, gold and crude oil moved to unprecedented highs, and the U.S. dollar declined erratically.

There are some interesting numbers in Table 15.11. For example, the S&P tends to open opposite to the previous weekly move; that is, if the Friday-to-Friday move was up, then there is only a 37% chance that Monday's opening will be higher than Friday's close. On the other hand, Monday has a 47% chance of closing in the same direction as the previous weekly move, so there is a good chance that buying a lower open will produce a profit by the end of the day. A similar pattern can be seen in the S&P for the other two trend conditions, and the same for U.S. 30-year bonds. Other markets did not show enough variation to be interesting.

Traders should consider that the three trend qualifiers only use one week of data. Trends in many markets last weeks and months; therefore, the same study could be done with a longer trend, which is more reflective of the positions held by investors.

The programs to create these tables, *TSM Weekend Patterns*, is available on the Companion Website.

Combining Weekend and Weekday Patterns

Taking the same seven markets, we can see if particular weekday patterns forecast the way prices will act on Monday. When doing this, it is important to remember that by increasing the number of cases from 3 (as in the trend shown in Table 15.11) to 16, the possibility of getting a good result by chance has increased.

In Table 15.12 the most interesting results are highlighted. The last two rows give the number of extreme probabilities for moves in the same direction and in the opposite direction as the previous Monday (always indicated by an X). Crude oil and the EURUSD show 10 combinations while most other markets have 9 and 8.

Looking at the EURUSD for the pattern XOXOO, there is a 69% probability of continuing on Monday's open. That is, if the previous Monday was higher than the Friday before, and Tuesday reversed, Wednesday was up and Thursday and Friday down, then there is a 69% chance that Monday will open higher. Out of 16 combinations for each market, about 9 showed combinations that were significantly far enough from 50%.

A Comment on Testing and Holidays

The weekday and weekend tests did not include any weeks in which there was a holiday or long holiday weekends. It would have been perfectly reasonable to view the opening of the first day of the following week as a "Monday" even though it occurred on a Tuesday after a Monday holiday. In fact, those special cases may have proved very interesting. Similarly, a holiday on Friday could have been worked out to fit the weekday patterns and allow the Monday direction, compared to the previous Thursday direction, to be added into the total results. These tests only used complete 5-day trading weeks.

A well-known work that studies price movement prior to holidays is by Merrill and can be found in Chapter 10 under the section, "The Holiday Effect for Stocks."

Day of the Month

Not all buyers and sellers try to time the market. The U.S. stock market has been used as a long-term investment for retirement portfolios for decades, and the amount of money being moved in and out of the market has increased significantly during the past 20 years. Institutions, such as Vanguard and Fidelity, that control IRAs, 401Ks, and other retirement plans have an obligation to put new money into the market, or redeem it, by the end of the month. Not all investments have daily liquidity. In addition, some firms must realize their gains or losses at the end of the month for accounting purposes. Non-U.S. firms, investing in the U.S. markets, may close out their profitable positions at the end of the month and repatriate their gains.

Does this show up as a pattern in price movement at the beginning or end of each month? To find out, we tabulated the returns (the daily percentage price changes based on continuous futures) from the first to the tenth day of the month, and from the last to the tenth-from-the-last day of the month. Using average returns for each year, we are able to see the current patterns and the changes in patterns.

Beginning with the S&P, Figure 15.22a shows the average returns for days of the month from 1982 through 2000. It is clear that there is a pattern to buying during the first three days of the month, mostly on the first day. Either that's when the investment house enters its inflows and/or a large number of investors set new positions based on reviewing the market, and their portfolios, at the end of each month. The end of the month has no outstanding pattern.

The past 10 years, 2001–2011, shows a very similar pattern (Figure 15.22b), with returns on the first day of the month very similar in size to those during 1982–2000. The end of the month shows slight buying, but not clear enough to draw conclusions. The period following the first two days of the month and preceding the last four days of the month is consistently lower. We can conclude that most buying takes place at the beginning and
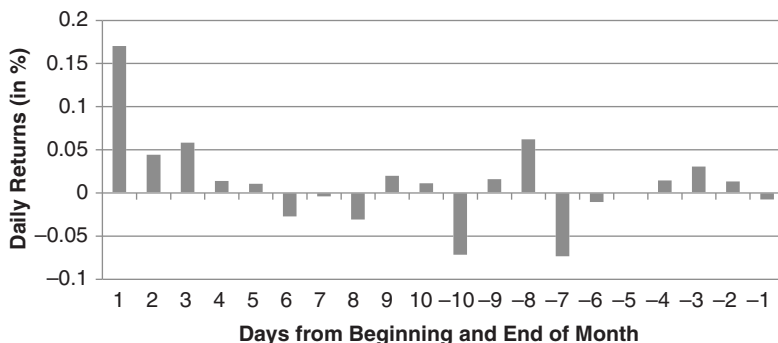
FIGURE 15.22a   *e*mini S&P futures day-of-the-month patterns, 1982–2000.
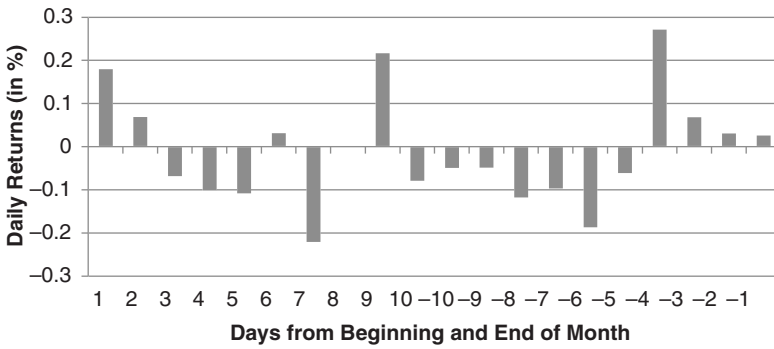
FIGURE 15.22b *e*mini S&P futures day-of-the-month patterns, 2001–2011.

end of the month, although counting on less volatile days during the middle of the month can always have an unpleasant surprise, as seen on day 9.

Eurodollar Interest Rates

Of course, there is much more money moved in the interest rate markets than in equities, and government debt is considered the most secure, even given recent uncertainty caused by the debt crisis. Looking first at Eurodollar futures, representing the short end of the yield curve, the 1981–2000 patterns, seen in Figure 15.23a, are erratic except for buying in the last three days of the month. Considering that those 20 years represented a bull market (rates declining from 18% in 1981), we might have expected more buying throughout the month.

The most recent 10 years, 2001–2011, is intuitively more reasonable, as seen in Figure 15.23b. First, most of the days are higher, confirming the upward trend of prices (downward yields). Also, the first five and last four days of the month show significant buying pressure, leaving only a few days mid-month for consistently lower moves.
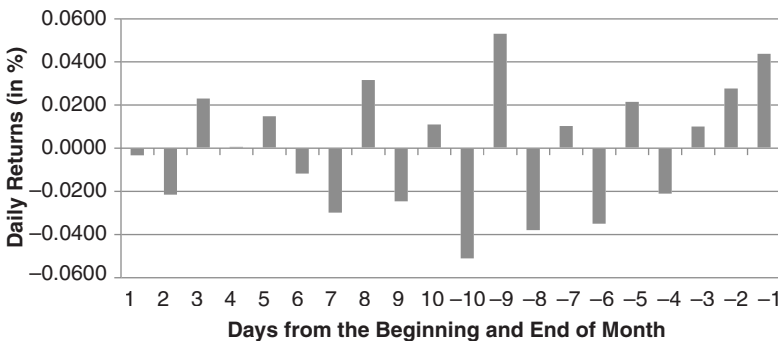
FIGURE 15.23a Eurodollar futures day-of-the-month patterns, 1981–2000.
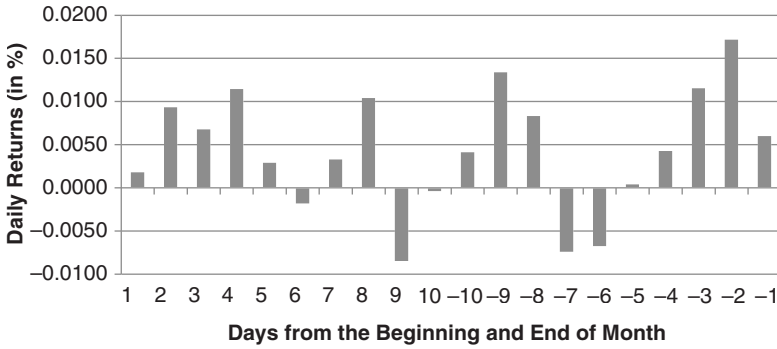
FIGURE 15.23b   Eurodollar futures day-of-the-month patterns, 2001–2011.

Crude Oil, Gold, and Soybeans

Do these patterns show up in commodities? While investors have used bond and equity markets in their retirement funds for many years, commodities had no role. However, in the past 10 years, commodity funds have become more popular, led by the Goldman Sachs Commodity Index (GSCI). You can buy into that index and benefit from effects of inflation on commodities. Figure 15.24 shows the GSCI Total Return Index (GSP) from July 2006 to July 2011, including the exceptional rally in crude oil and grains in the summer of 2008 followed by an equally exceptional reversal. Since then, prices have again gained 50%.

Crude oil, Figures 15.25a and b, shows the same patterns as the other markets. During the past 10 years, the last four days of the month are strong, with the last day the strongest. This is in contrast to the first 18 years, when the last day of the month was sharply lower. Even though we think about the last 10 years of oil as moving higher, there are some large downward moves during the month.
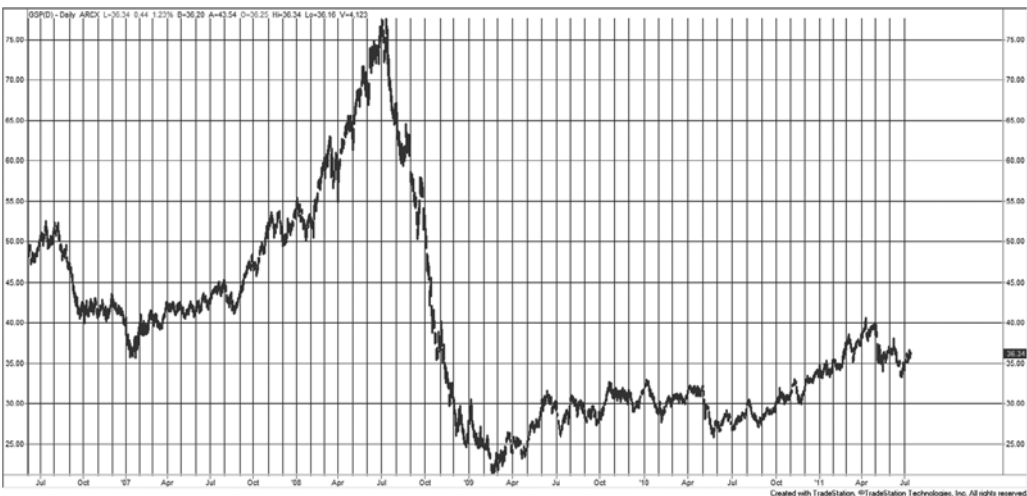
FIGURE 15.24   The GSCI Total Return Index (GSP) from July 2006 to July 2011.

Gold has a more pronounced change than crude, with the end-of-month showing lower moves for 1975–2000 and higher for 2001–2011 (see Figures 15.26a and b). While this seems reasonable, be careful to understand that gold declined from its highs in 1980 and kept declining into the early 2000s. Its rally has only occurred in the past few years. Given that, it is still important that buying is consistent during the last six days of the month when it was the opposite in previous years.

And then there is soybeans, a very important U.S. crop, although more attention seems to go toward wheat, corn, and cotton. When any one of those markets rallies, they all rally because they are mostly affected by the same weather conditions, and the grains are feed substitutes for one another based on their protein ratios. The grains have the longest trading history of all futures markets, going back to the first part of the twentieth century, longer if you search harder. Figures 15.27a and b show the change in the soybean pattern. From 1949 through 2000, the beginning of the month is relatively quiet compared to the end of the month, but the end shows significantly more declines, with an upward move limited to the last day.

Days-of-the-Month Summary

Average results do not mean that all months will perform as you wish. Some could be a substantial reversal even while the averages are positive. What seems clear across all markets is that the past 10 years have seen a shift toward buying in the last few days of the month. And, except for Eurodollars, there is also a pattern of selling during the first few days of the month, and buying on the first day of the next month, followed by selling.

How can you use this? If you need to buy for a long-term position, then buying about four days before the end of the month would take advantage of what is called *free exposure*, that extra boost from others buying after you. It may be possible to construct a strategy that buys four days before the end of the month and closes out that position on the end of the first day of the month.

These patterns may change if there was a strong trend. Gold and crude oil may be the best examples. If the economy goes into a sustained contraction, declines may outnumber rallies, and the buying at the end of the month would become less significant while the selling after the first of the month might increase dramatically.

Reversal Patterns

The last of the three studies looked at reversal patterns. The intention was to find a pattern in the open, high, low, and closing prices of the day that would help predict the next day's pattern or direction. Three combinations are:

1. *Intraday trend continued.* An *upwards trend* is set up when yesterday's high is greater than the previous high and yesterday's close is greater than the previous close. We then find the percentage of days that open higher or close higher than yesterday. A *continued downwards trend* is the opposite pattern.
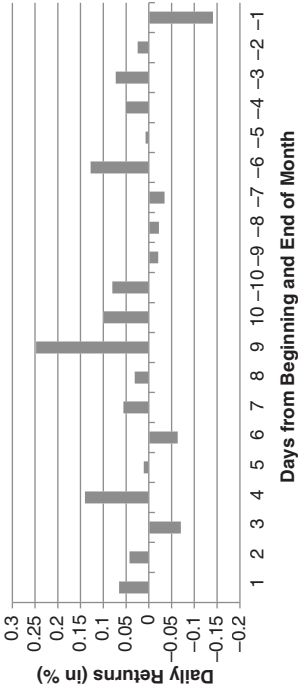
FIGURE 15.25a Crude oil day-of-the-month patterns, 1983–2000.

FIGURE 15.25b Crude oil day-of-the-month patterns, 2001–2011.

FIGURE 15.26a Gold day-of-the-month patterns, 1975–2000.
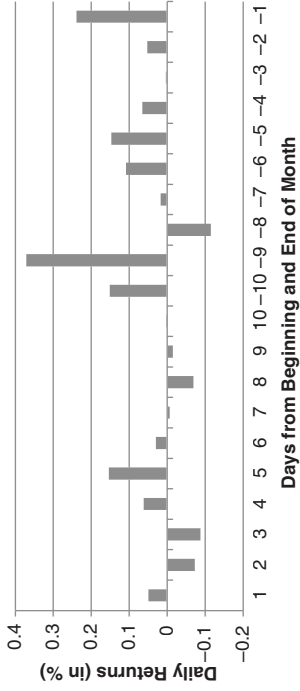
FIGURE 15.26b Gold day-of-the-month patterns, 2001–2011.

FIGURE 15.27a Soybeans day-of-the-month patterns, 1949–2000.
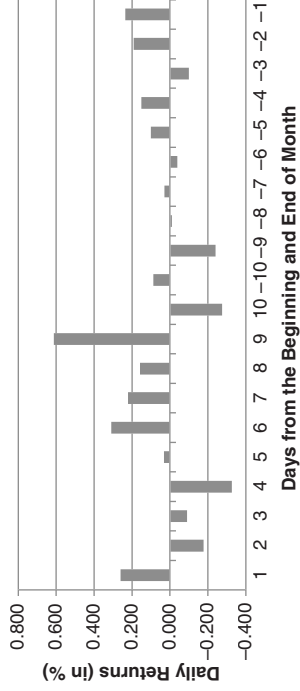
FIGURE 15.27b Soybeans day-of-the-month patterns, 2001–2011.

730

2. *Reversal day.* A very popular chart pattern called a *downward reversal day* posts a higher high than the previous day but a lower close. Most traders believe that there is follow-through into the next trading day. An upwards reversal day has a lower open and higher close.

3. *Outside reversal.* This is a more extreme case of (2), where yesterday's high was greater than the prior high, but yesterday's close was below the prior low. This downward outside reversal is continued if the next day opens lower or closes lower.

Results for these trend and reversal combinations are shown in Table 15.13 for the eight markets representing a sample of futures and stocks. The trend is the least distinctive pattern, with the exceptions of the S&P and wheat. When the trend day is higher, the next day does not show much of a bias, but a small tendency to open lower and close higher than it opened. For downward trending days, all markets are more consistent about opening lower and closing lower. The S&P, however, has a strong tendency to open

TABLE 15.13 Reversal Patterns, January 2000–July 2011

		Higher High		Lower Low		Approximately 2800 Cases	
	Pattern	Open	Close	Open	Close	Up	Down
EURUSD	Trend continued	53	49	49	47	1028	910
	Reversal day	48	45	55	51	403	395
	Outside reversal	54	43	53	50	97	94
10-Year Notes	Trend continued	50	52	50	45	1053	871
	Reversal day	47	46	58	57	335	331
	Outside reversal	45	44	58	53	77	60
Gold	Trend continued	48	50	49	45	1061	890
	Reversal day	53	44	44	45	301	248
	Outside reversal	56	41	37	43	66	63
Crude	Trend continued	48	50	48	46	1044	906
	Reversal day	51	45	57	49	380	330
	Outside reversal	49	47	49	44	90	68
Wheat	Trend continued	46	46	32	48	868	992
	Reversal day	43	50	54	42	388	357
	Outside reversal	34	41	44	43	79	72
S&P	Trend continued	26	52	41	43	1022	890
	Reversal day	52	43	40	53	421	425
	Outside reversal	48	40	39	56	115	88
AMZN	Trend continued	43	49	48	48	997	956
	Reversal day	50	47	47	46	367	365
	Outside reversal	57	46	46	41	65	80
XOM	Trend continued	46	48	41	44	1089	937
	Reversal day	45	39	45	50	400	330
	Outside reversal	41	41	39	50	80	66

lower after an upward trending day, then close neutral (equally up or down). Wheat is the opposite, having a strong tendency to open higher (because there is a 32% chance of it opening lower) after a downtrending day, and an equal chance closing either higher or lower. These markets both present significant opportunities. The S&P also acts as though it has significant noise, which was discussed in Chapter 2.

Reversal days do not occur as often as trend days, as seen in the two right columns; however, the results are interesting. Numbers less than 50 show that there is no follow-through of the pattern into the next day. For example, the S&P shows extreme results. For days with higher highs and lower closes, the next open has a 52% chance of continuing lower, but the close of the following day only has a 43% chance of being lower; therefore, the opening reversal is not sustained, and prices tend to rally from the open to the close. For a lower low and a higher close (an upwards reversal), there is the opposite pattern, only a 40% chance of a higher open but a 53% chance of a higher close. We then get a rally from open to close with either pattern. This could confirm the overall upwards bias in the index markets, but these numbers are larger than the expected bias. Crude oil also has an interesting pattern, the same as the S&P for downward reversals, but the opposite for the upwards reversals.

The outside reversal day, which is a much smaller number of events, is most interesting for the S&P. For the higher high and close below the low, there is very little follow-through on the open, but a rally by the close. For the lower low and close above the high, the next open is lower, but the close is higher. Then the S&P exhibits noise for the next open but is biased toward an upwards move on the close. Again, this upwards move seems to be greater than the general upwards bias attributed to the S&P and other equity index markets.

The program, *TSM Reversal Patterns*, is available on the Companion Website and produces a file in the same format at Table 15.13.

COMPUTER-BASED PATTERN RECOGNITION

The methods previously discussed were based on patterns familiar to traders. The weekly and weekend studies, as well as the intraday time patterns shown in the earlier sections, were verified by hand in the first edition of this book, but eventually succeeded in becoming one of the earliest computer studies. Now, of course, we can test many patterns across many markets looking for both unusual situations and similar occurrences. There is a type of pattern recognition, however, that would never have been done without computing power.

Rather than the conventional price patterns where recurring sequences of higher and lower days are found within certain qualified intervals, *computer-based pattern recognition* refers to *sets of descriptors* and *classes of interest.* For example,[6] Aronson

---

[6] David R. Aronson, *Artificial Intelligence Methods* (privately published). Also see David R. Aronson, "Artificial Intelligence/Pattern Recognition Applied to Forecasting Financial Market Trends," *Market Technicians Association Journal* (May 1985): 91–131.

describes the sets and values which must be satisfied by a professional jockey as follows:

Descriptor	Value Range
Height	Under 5′ 5″
Weight	Under 120 lbs.
Age	16 to 35
Years riding horses	Over 10

The *set* of people who satisfy all four conditions are said to *contain* all professional jockeys. The converse, that all people satisfying these conditions *are* professional jockeys, is not true. It will be necessary to qualify this set further to create a set that contains *only* professional jockeys; however, these four conditions go a long way toward reducing the field.

How can this tool be applied toward the development of a trading strategy? If the system is defined in terms of the *trade profile*, it becomes obvious. Consider the following characteristics of a trade:

1. The price moves higher or lower by at least 3% of the starting price.
2. The price move occurs within 20 days.
3. There is no loss exceeding 0.5%.

Either a computer or an analyst can locate all price moves that satisfy these conditions within a price series. Each of the 5 days preceding an upwards move, which satisfies these conditions, can be marked as *buy* days, and the 5 days preceding a downwards movement can be designated *sell* days (see Figure 15.28).

FIGURE 15.28   Specific buying and selling days.

By adding to a table as these combinations are located, the computer now contains the set of all buy and sell days; those days on which it would be good to get a buy or sell signal based on historic data. Next, some likely indicators must be specified to identify when these trades are about to begin. The following might be used:

- The moving average direction
- An overbought/oversold indicator such as the *Relative Strength Indicator* (RSI) or *contrary opinion*
- The direction of changes in trading volume
- A 10-day momentum indicator

By entering a broad selection of indicators and trying to avoid duplication, the computer can find unique values for combinations of indicators that primarily occur during the days selected as buy and sell periods. Ideally, all buy signals should occur when one indicator, or the value of combined indicators, exceeds a specific threshold. For example, all buy signals occur when the average value of the RSI and the market sentiment (contrary opinion) is below the 30% level; however, having all buy signals occur here is not enough. Poor signals may appear at this level, which cause large losses. The perfect system will have no losing signals occur in this zone.

Unfortunately, in the real world there are no perfect solutions. The trades that are signaled by the combinations of indicators will have to be studied for net return, risk, and other performance criteria. However, the technique of setting up classes of indicators, and buy and sell days, is a new and valid approach to system development. It is analogous to the multiple regression method used by econometricians to find the relationship between statistics and prices. Although the econometricians use inflation, supply, interest rates, and so forth, pattern recognition can employ technical indicators and discrete price patterns to forecast a buy or sell day.

A computer program designed to identify what conditions are common to a selection of price moves is available in LIM (Logical Information Machines), a service that provides an exceptional amount of economic data as well as a querying language in which you can specify problems such as this type of pattern recognition. When you hear the financial news broadcasters saying, "today's drop of 3% in NASDAQ has a 90% chance of reversing tomorrow," that information could be extracted through LIM, and could be either a broad statement or one conditioned on a reaction to a particular economic report.

Analysts interested in these techniques will find neural networks and genetic algorithms of particular importance. They are powerful search tools that find patterns in a large number of inputs with remarkable efficiency. Detailed discussion of these methods can be found in Chapters 20 and 24.

Repeated Patterns

A mindless task for a computer is to recognize repeated price patterns. Over many years, the chance of a higher close following a higher close (or even a lower close) in the stock

market is about 55%, due to investor bias and the uptick rule. In most other financial and commodities markets, there seems to be very close to a 50% chance of any day being higher or lower. If there were a 65% chance that a higher day followed a lower day, traders would jump on that opportunity and extract all possible profits until that pattern was extinct.

What about two higher days in a row, or two lower days? We saw something similar in the Weekday Study, where various combinations that occurred during the week were posted along with the frequency of a similar or reversal day following. There were some combinations that showed high or low frequencies and attracted our attention. In the Weekday Study we limited the analysis to repeated days of the week; however, these patterns could start at any time.

When seeking repeated patterns, it is important to remember that the number of cases drops quickly as the length of the pattern is extended. For example, if you are testing 1,000 days of data looking for a 1-day pattern, you have 1,000 candidates. About one half will be higher days and the other lower days. If you now look for 2-day patterns, for example, up-up, then only 25% of the days are likely to qualify, because the other three combinations, up-down, down-up, and down-down, will normally share equal amounts. With 3-day patterns there are 9 combinations, each with about 1/9th of the data, and so on. By the time you are looking for a sequence of 10, there are 100 different combinations and the number of occurrences for each pattern will be very small.

But what if a specific sequence of 10 ups and downs can be found 10 times in a series of 1,000 prices? And what if an up day followed that sequence 70% of the time? Would you have a potential trading system? Unfortunately, there are just not enough cases to have confidence in the result.

One way to improve the sample is to look at the total count of up and down days within a long sequence, rather than the exact pattern of ups and downs. Therefore, if you found a sequence of 20 days with 15 ups and 5 downs, you would have a very bullish scenario for nearly all combinations, the least bullish being 5 down days at the end. You have now greatly expanded the number of patterns that satisfy this condition, and you can give this sequence a quality—bullishness—which may improve the chances for predictive success.

Readers interested in this method should read the sections on "Gambling Techniques—The Theory of Runs" and "Martingales and Anti-Martingales" in Chapter 22.

ARTIFICIAL INTELLIGENCE METHODS

Artificial intelligence refers to a computer process that performs an operation corresponding to or approaching human thinking. This is intended to distinguish it from simple pattern recognition, with which it is often confused. The state-of-the-art in artificial intelligence is the separation of two ideas. The collection of information that is stored in the brain has been termed the knowledge base. This is distinguished from reason, rules, and logic, called the inference engine. These ideas are not very different from the database and trading strategies that are discussed here.

The closest practical approach to artificial intelligence is heuristic programming. This refers to computer learning in very much the same way as a laboratory animal finding its way out of a maze. The computer starts with rules relative to the problem, then records the successful and unsuccessful experiences. Eventually, it has a complete table of what to do for each situation or at least a table of probable solutions. This is a realistic, intelligent approach when the same events can be expected to recur in the same way. It does not help in new situations without the added complication of extrapolation, basic relationships (e.g., price level to volatility), and other forms of expectation. This separates the problem of trading the market successfully from winning a chess game, in which there are a limited (although large) number of combinations.

The danger of the heuristic approach to pattern recognition is that it may continue to define longer combinations of patterns that have already produced inconsistent or poor results. Allowing the computer to identify a limitless collection of patterns is just another case of overfitting, but this time at a highly sophisticated level.

Heuristic programs have improved current technology in searching, optimization, and game-playing strategies; however, they are not readily available to the trader. There is no doubt that this technique will be quickly absorbed into trading strategies as they become available.[7] Those readers interested in these methods should learn about expert systems, neural networks, and genetic algorithms, which can be found in Chapters 20 and Chapter 24.

---

[7] A book that covers the basics of heuristics is Judea Pearl, *Heuristics* (Reading, MA: Addison-Wesley, 1984) and is still available. Recent books on computer game playing focus on *role playing* rather than strictly games of chance; however, a book by David Levy, *Computer Chess Compendium* (reprinted by Ishi Press, New York, 2009) was followed by David Levy and Monty Newborn, *How Computers Play Chess* (New York: Computer Science Press, 1991), and should give insight into tactical solutions.

Day Trading

A *day trade* is a position entered and liquidated during the same trading day. Day trading reached a peak in popularity along with the top in the stock market at the beginning of 2000. High volatility and a clear bull market made day trading look easy. The *opening range breakout*, buying a new high after the open and selling on the close, was the *strategy du jour*, but that period did not last long. While trading for a living is always difficult, day trading puts more restrictions on the rules, requiring you to extract a profit during a very short time period. Intraday prices respond more to noise than to fundamental factors. They jump up and down following economic reports, large orders from funds, earnings statements, upgrades and downgrades, and the flow of microanalysis on the financial news networks. The fact that the economy is strong or weak, that creeping inflation requires raising interest rates, or that the budget deficit is steadily growing has little impact on a trade that targets a 1-hour period during the day. Day trading is highly focused on price patterns, noise, and volatility.

Gaps, time of day, and various daily patterns discussed in the previous chapter (time patterns can be found in Chapter 15) are easily applied to day trading. Day traders often look for extreme price moves and a predictable pattern that follows. Day trading requires exceptional discipline, excellent planning, anticipation, and concentration. The need for a fast response to changing situations tends to exaggerate any bad trading habits; as in other fields, the shorter the response time, the greater the chance for error. In this chapter we will extend the idea of day trading to systems and methods that may also be held overnight, but expect to limit the trade to about 24 hours.

Day trading has become more sophisticated as computers, electronic exchanges, direct access to live data feeds, better graphics, and more programmable strategies have descended on the industry. There is so much competition trading into NASDAQ that day traders target as little as 1¢ per share net profit on a trade. This is only possible because commission costs have dropped dramatically for electronic execution. The majority of exchanges in the world are now electronic, and any new markets or exchanges will be electronic. GLOBEX, the electronic offshoot of the Chicago Mercantile Exchange, allows side-by-side, 24-hour trading of a number of markets, most notably the S&P 500; however, pit trading is rapidly declining.

EUREX, the largest European exchange, has been fully electronic since its inception; France's MATIF, London's LIFFE, and Sydney's SFE have all moved from pit trading to electronic markets. This is all good news for the traders. In addition to faster executions, the data streams have fewer errors, and there is little question of a fair execution price. These changes have greatly increased competition among individual and commercial traders. Market makers have been squeezed out by penny price changes (compared to ⅛ cent) and high frequency trading.

For the arbitrageur, computers have had an even stronger impact. Sophisticated systems at banks and large financial institutions consolidate data feeds that bring current transactions on every type of interest rate vehicle in every maturity and major currency. Analytic programs can find outliers and show which combinations (such as *strips*) can produce a "riskless" profit. For the individual trader, few of these opportunities are available, although trading by these commercials adds liquidity to the market. Individuals, however, find it much easier to create spreads of different deliveries within the same market as well as spreads between two related products. Stock traders can create sector baskets or look for performance differences within a sector. Spread trading opportunities, based on formulated values such as the energy crack or soybean crush, can also be recognized and signaled automatically. Many of the opportunities that now seem so easy to see would previously have been missed. This faster, more systematic response to the market allows traders to improve profits and reduce risk in any day trading method. The independent trader, willing to take a larger risk and hold positions longer, still has great opportunity.

IMPACT OF TRANSACTION COSTS

Transaction costs are the greatest deterrent to day-trading profits. The two components of costs are the fees paid to the broker and the *slippage*, the difference between your execution price (the real filled order) and your intended price. The combination can remove a large part of potential profitability and even turn expected profits into unexpected losses.

The fixed costs of trading have plummeted over the past 20 years. As far back as the early 1970s you would have paid as much as 1% of the share value to trade stocks, and $50 per contract (each way) for futures. Stocks are now priced at $7 to $10 per transaction, where a transaction can be thousands of shares. The larger the trade, the cheaper the per share cost. Futures are more expensive per contract, ranging from $3 to $15, but contracts are far bigger than a typical stock trade. One contract of gold is 100 ounces, at a value of $1500 per ounce is $150,000. Then for $10 you can buy or sell $150,000 worth of gold. For the position trader, one who holds the trade for weeks, the cost is incidental. For the day trader, it needs to be watched, as do all costs.

Costs and Volatility

An aspiring day trader has three ways of improving performance: paying lower commissions, carefully selecting opportunities, and executing better. Part of selecting carefully is choosing markets with higher volatility. If volatility is greater, that can translate into larger profits, which then make costs a smaller percentage. Table 16.1 shows the average

TABLE 16.1 Volatility and Volume of Selected Markets, 1995–2010

Futures

$Volatility	1995	2000	2005	2010
EURUSD	2016	1436	1413	1937
Crude	307	998	1692	2186
S&P emini	224	1478	600	931
DAX	665	3845	1284	2674
Gold	222	330	522	1981
10-Yr Note	553	568	526	777
Wheat	414	279	371	1008

$Volatility	1995	2000	2005	2010
EURUSD	1.08	1.24	0.89	1.17
Crude	0.59	1.25	1.42	2.42
S&P emini	0.49	1.85	0.93	1.66
DAX	0.61	1.76	0.92	1.71
Gold	0.35	0.71	0.90	1.60
10-Yr Note	1.03	0.95	0.57	0.66
Wheat	0.58	0.51	0.83	2.79

Average Volume	1995	2000	2005	2010
EURUSD	29,424	16,999	138,115	340,837
Crude	94,035	147,616	237,296	669,270
S&P emini	74,675	76,236	821,804	2,203,560
DAX	39,016	45,357	126,386	159,770
Gold	31,055	26,049	63,510	176,818
10-Yr Note	89,788	185,998	814,723	1,160,889
Wheat	19,663	25,428	39,666	91,427

Stocks

$Volatility	1995	2000	2005	2010
BA	0.65	1.85	1.16	1.72
BAC	0.27	0.97	0.59	0.46
GE	0.17	1.73	0.45	0.42
AMZN		4.24	1.1	4.13
XOM	0.24	1.17	1.19	1.06
MRK	0.44	2.25	0.63	0.73
GOOG			7.27	11.48

$Volatility	1995	2000	2005	2010
BA	2.15	3.82	1.86	2.60
BAC	1.84	3.99	1.31	3.16
GE	1.70	3.33	1.28	2.52
AMZN		9.12	2.75	2.97
XOM	1.38	2.82	2.04	1.63
MRK	1.87	3.22	2.04	2.02
GOOG			2.62	2.14

Average Volume	1995	2000	2005	2010
BA	2,282,684	4,069,678	3,890,172	5,850,680
BAC	2,905,781	11,573,837	11,295,613	188,980,271
GE	12,222,984	16,425,906	20,033,315	72,251,080
AMZN		8,696,973	7,210,178	7,294,318
XOM	5,015,275	10,893,183	18,896,922	26,463,757
MRK	5,469,516	6,060,400	10,029,535	15,242,295
GOOG			10,733,174	3,182,011

daily price moves in dollars or euros, the average percentage volatility, and the average volume, for a selection of futures markets and stocks over intervals of five years. The average volatility is based on the *true range*, which includes the previous day's closing price. For those who are strictly day traders, the actual high-low range is better; however, the result should be the same for stocks and only slightly lower for futures.

For futures markets, some of these numbers make more sense as percentages, and others as absolute dollar-value moves. For example, commodities (crude oil, gold, and wheat) show exceptionally large increases in dollar value as prices have doubled and tripled over the years. Shown as a percentage of the average yearly price, that increased volatility is not as extreme.

Interest rates are quite different. The percentage is also calculated on the average price, but that price would be better as a yield. Because interest rate prices are quoted opposite to yields, yields go down as prices go up. Typically, volatility increases with higher yields, so that percentages based on price should be viewed skeptically. The dollar volatility of T-Notes has stayed relatively constant, jumping up a bit in 2010, but the percentage volatility has declined steadily as prices have kept rising—the divisor has gotten consistently larger while the yields have fallen. The EURUSD has a similar problem because, as the U.S. dollar declines, the value of the euro becomes a larger denominator.

Volume, however, is clear. For both futures and stocks, volume has increased many times since 1995.

There are always exceptions, and although prices have dropped since 2007, uncertainty has kept volatility higher than normal. A comparison of the average volatility of the stocks in Table 16.1 with the *e*mini S&P shows that our selection of stocks has proved to be unusually volatile because they are companies that have been in the news over the years and more actively traded.

	1995	2000	2005	2010
Selected stocks	1.79	4.39	1.99	2.43
S&P futures	0.49	1.85	0.93	1.66

If you are strictly a day trader, you should also know that the range from the open to close for the S&P is only ½ the total range for the day. If you are planning to enter on the open and exit on the close, you have already severely reduced your opportunity. However, knowing that, entering at better price than the open is a good goal.

Costs and Trading Strategy

The cost of trading is different if you buy and sell in the direction of the price move (*trend trading*) or if you are trading against the direction of the move (*mean reverting*). Day traders can be either one. Trend traders are likely to have higher costs, entering new long positions as prices move higher and selling short on new lows. For example, in trading the *e*mini S&P, worth $50 per point, slippage should be expected to be at least ½ point, or $25, for each entry and exit. If the market is volatile, or moving fast, slippage can be much more.

If the SP daily range is $931, the fixed cost per side is $10, and the slippage is $100, then there is $821 remaining. If you could capture half of the trading range (the same as the open-to-close range), or even one-quarter of the range, there would still be sufficient net profits, provided a large number of your trades were profitable. Typically, trend trading has only a ⅓ chance of a profitable trade, and the profitable trades are twice as large as the losers, then your net result is a loss of the total transaction costs (one profit at $90 for two losses each at $45 is a break-even). You will need to do better than the average trend system to survive.

One price characteristic that is exploited in day trading is market noise. As discussed in Chapter 2, equity markets have the greatest amount of noise, and this noise can be used for timing. Rather than chasing an entry price, waiting for prices to reverse will reduce risk and increase the chance of a profit.

Liquidity

With the *e*mini S&P daily volume twice 10-year Notes and nearly four times crude oil, it is no surprise it is the preferred choice of day traders. High volume translates into liquidity and makes slippage relatively small. While the selection of day-trading candidates begins with those markets of greater volatility, volume is very important. Whether you are trading 1 contract or 1,000, a thinly traded market produces slippage that will cut sharply into profits. In choosing among index markets that usually offer the greatest volatility and profit potential, you are always safest with the markets that combine the highest volume and highest volatility.

Markets with light volume can show larger price moves than similar markets traded on other exchanges. Traders will be tempted to profit from these moves but will consistently find that the execution of an order (either buying or selling) at the posted price is elusive. While you see the price on the screen, the bid-asked can be very different. The reality of trading these markets is that a Market Order or a Stop Order is not advisable due to the thinness of the trading. A Limit Order may not get filled, and a spread is not quoted at anything resembling the apparent price relationships that you see on a quote screen; there is no real way to take advantage of these perceived profits. If an execution succeeds, exiting the position still has the same problems plus some added urgency.

Mean-Reverting Strategies

Some strategies have less slippage than others. While trend following and momentum techniques will buy and sell in the direction of the price move, mean-reverting strategies do just the opposite. They assume that relatively large price moves will not continue; therefore, a sharply higher move is *sold*, and a sharply lower move is *bought*. Trading against the direction of the market with a small order should allow an entry at the intended price, reducing the slippage to near zero. Exiting a mean-reverting method may not be as easy. If successful, prices reverse, and you can exit at your leisure when prices return to a normal point, where they can remain for some minutes. However, if prices start higher and continue higher after you have sold, at some point

you will want to cut your losses by buying back your shorts. This means buying when prices are moving higher and is subject to the same slippage as a directional strategy. The impact of slippage on a mean-reverting strategy depends on how often it achieves its profit targets because taking profits is also trading against the direction of price movement. Conservatively, you can assume that a mean reversion program has only 50% of the slippage of a trend program. As an added bonus, stock commissions may be significantly lower when a sell order is placed above the market, or a buy order below the market, because they serve the function of adding liquidity. Market noise and lower cost of slippage are primary reasons why many day traders choose mean-reverting strategies.

Missed Orders

With a systematic day trading program, your success depends on executing *all* of the signals reasonably close to your system price. During testing, you should have accounted for reasonable slippage and commissions in the net results. In some cases that is not enough to ensure success. If you are buying and selling in the direction of the price move, there are days when the market jumps after a news release, and you cannot get filled anywhere close to your intended price. With most of the potential profit gone before you enter the order, it would seem reasonable to simply skip that trade. These missed orders, called *unables*, can add up to a large part of your profits; at the same time they never reduce your losses.

Some markets are prone to more *unables*. Those that are open when the morning economic reports are released are most often a problem. The interest rate futures open 10 minutes before the regularly scheduled government economic releases and are often very quiet at the open while traders wait for the reports to be posted. European markets, both interest rates and equities, as well as all foreign exchange markets, are actively trading in advance of these reports and can change direction abruptly when the news is a surprise. On Wednesdays, the American Petroleum Institute (API) releases energy stocks at 10:30 A.M., Eastern Time, 30 minutes after the energy markets open for trading. On Thursday, natural gas stocks (AGA) are reported at 10:30 A.M., also 30 minutes after its open. Other reports, such as the Chicago Purchasing Managers and Michigan Sentiment Index are typically released at 9:45 and 10:00 A.M., Eastern Time, after the stock market has opened. The Fed announces the results of the Fed Open Market Committee (FOMC) meeting at 2:15 P.M., Eastern Time, 15 minutes after the interest rate floor session ends. Add unexpected political and weather news to these regular announcements, and you can easily form the opinion that prices are driven by a series of price shocks. The larger jumps make it difficult to execute a day trade in the direction of the price move, except at the worst price of the move.

For the energy complex, a flare-up of military or political activity in the Middle East has caused a prolonged period of very erratic price movement resulting in as much as 20% unexecuted trades during a 1-month period using a directional strategy. If we consider the normal profile of a short-term trading system as having an average net profit of $250, an average loss of $150 and a 50% frequency of profits, we expect a profit of $5,000

for every 100 trades and a reasonable profit-to-loss ratio of 1.66. If, however, there are 10% *unables*, those missed opportunities must come from the profits; if the market was moving in the opposite way, you would get all of your positions filled. Then 10% of the profitable trades means 5 trades missed out of every 100 for a total of $1,250. This reduces the total profits to $4,750 and the profit factor to 1.50. It may turn a marginal trading strategy, or one with small profits per trade, from a profit to a loss.

Markets with Trading Limits

Day traders may find that those markets with traditional daily trading limits present a problem during high volatility periods. These are mostly agricultural markets, although some financial markets have a pause in trading at trigger levels. In April 2011, the CME raised the daily price limits in corn to $0.50/bushel. Previously they had a process of expanding the limits on each day following a limit move.

Day trading does best in markets that have wide swings not deterred by limits; a single locked-limit move can generate a loss that offsets many profitable day trades. High volatility and locked-limit moves present a contradiction for day trading. Expanding limits and trigger levels have greatly helped reduce the frequency of locked-limit days.

Estimating Slippage Costs

If you know the cost of slippage, you can do a much better job selecting the markets to trade and have a realistic appraisal of your trading expectations. The factors that make up slippage are volatility, overall market volume, the size of the order being placed, and specific market activity. Of these items, the specific market activity is the most difficult to record, because it can have any combination of noise and directional movement.

Total intraday volume is now available on Bloomberg for many markets, although the exact size and depth of orders is available only on more sophisticated trading platforms. Being able to *hit the bid* means that you get exactly the price you see, but it does not mean that it is the system price, so there will still be some slippage. Some data services provide *tick volume*, the number of price changes (actual trades or "deals done") during the intraday bar. In general, tick volume is directly proportional to actual trading volume. During periods of greater activity, both contract or share volume will increase along with the number of price changes. It will allow a rule that says, for example, *buy on increasing volume*. This will be discussed at the end of this chapter.

If you have recorded the order price, execution price, volatility (daily difference in the high and low), daily volume, time of day, and tick volume, you can find the importance of each factor and estimate the slippage for any trade by applying the formula:

$$\text{Actual slippage at time } T = a_0 + a_1 \times \text{Volatility} + a_2 \times \text{Daily volume}$$
$$+ a_3 \times \text{Current activity at time } T + a_4 \times \text{Size of order}$$

By creating a spreadsheet of values for all trades, you can solve for a_0, a_1, a_2, a_3, and a_4, and then estimate future slippage at any time T during the day, given the current volume

and order size. This approach was taken by Greer and Brorsen[1] for large stop orders using a ratio of order size to current market activity. The results show that slippage was very modest for most markets.

For exiting on the close, day traders can use the same method as large macrotrend traders, placing an order that guarantees the settlement price. Similar to a *Market on Close* order, which only guarantees that the fill is within the last 30 seconds of trading, this *Trade on Settle* (TAS) order assures that you get the settlement price, the price used to mark-to-market each day.

Greater volume, as seen in Table 16.1, combined with electronic exchanges and order entry systems, has reduced slippage considerably. It is reasonable to expect average slippage to be less than 1 tick, that is, one minimum point move, for most markets. There are times, immediately after a surprising economic report, when prices jump and any order can be filled at a bad price.

Volatility-Adjusted Slippage

If exact numbers are not available, and you are trading during the day rather than on the close, it is reasonable to assume that slippage will be larger if volatility is greater. Whatever slippage you have decided to use, based on average price movement, you could increase that amount in proportion to increased daily volatility. However, decreasing slippage based on lower volatility should be capped, so that there is always some minimum amount of slippage. It is best to be conservative.

KEY ELEMENTS OF DAY TRADING

Systematic day trading has evolved along with the markets. During the 1990s, the opening range breakout produced great success by entering on a new high or low 15, 30, or 60 minutes after the open. Or, you could fade a large move in U.S. bonds after an economic report, and count on price reversing. In the early 2000s, pairs trading in stocks generated steady profits, as did a larger market neutral portfolio of pairs. Now it is high frequency trading, similar to pairs trading and market making, but done in microseconds (perhaps even milliseconds). The evolution has been from outright directional positions to complex mean reversion and spreads.

New tools have made it easier for any trader to participate. TradeStation, the most popular programmable platform, provides a way to enter your rules, then automatically send orders to be executed as the system generates buys and sells. Of course, not everyone is comfortable with letting a program run throughout the night, sending off orders on its own.

---

[1]Thomas V. Greer, B. Wade Brorsen, and Shi-Miin Liu, "Slippage Costs of a Large Technical Trader," *Technical Analysis of Stocks & Commodities* (January 1992).

Migration of Trading Techniques

Most traditional methods of technical analysis, in particular trend-following systems, are not used for trading when the holding period is less than 3 days, and probably do not apply for periods under 20 days. An economic analysis of supply and demand, or the effects of the government's interest rate policy, would also not be important to a day trader. However, trading the price reaction to economic reports can become a day-trading strategy, that is, trading the market after the initial price jump. While no one has yet figured out how to anticipate a price shock, the day trader will *react* to large moves, not anticipate (this is covered in more detail in the section "Identifying Price Shocks" in Chapter 22). Techniques used for day trading can be applied to slightly longer periods of a few days.

Short-term price movement is mostly noise. The part that is not noise is not particularly important although it may cause grief from time to time. Recognizing the extent of the noise directs a strategy in two ways:

1. For directional entries, waiting for a pullback will improve entries.
2. Mean reversion becomes the preferred strategy, because price direction is not sustained, whether up or down.

If you accept the premise that short-term movement cannot be distinguished from noise, then classic charting patterns, such as triangles and flags, are simply random patterns with no significance. Classic chart patterns that are limited to intraday data would not be a good basis for a strategy. However, there are very popular intraday patterns that will be discussed, where penetration of previous key levels is important as well as changes in volatility and volume.

From the floor trader's viewpoint, only the most obvious recent key levels are important. Today's price movements are compared to today's opening price; today's high and low, yesterday's closing price; yesterday's high and low, and less important, last week's high and low; very memorable older support and resistance points; and finally, the life of contract high and low. When the same price satisfies more than one condition, there is greater confidence in the importance of that point.

Point-and-Figure for Short-Term Trading

Point-and-figure charting had been a primary tool of day traders for many years. Using the minimum price movement as the box size and a 3-box reversal, many traders have kept a continuous, although lengthy, chart of day-to-day price movement. Buy and sell signals can be taken in the standard manner, but day traders are most likely to use these charts for identifying countertrend support and resistance levels. Intraday point-and-figure, as well as moving averages, often shows frequent changes of direction before a new buy or sell signal occurs, even when the minimum box size is used. Trading in the direction of the price move is best applied to short-term overnight positions where the size of the move is much larger. A primary advantage of point-and-figure is that directional signals occur concurrently with a price breakout. There is no lag when trading using point-and-figure.

Price reversals that occur during the day can be plotted using a smaller point-and-figure box size than is common for daily charting. The more frequent reversals due to the combination of more data and smaller box size will define the trend sooner or bring the protective stop closer. Remember, stop-loss levels advance only when there are reversals on the point-and-figure chart. Traders who use daily charts cannot see the different patterns that develop during the day.

In Chapter 5, where point-and-figure rules are given, there was a choice of plotting the reversal or trend continuation on an outside day. For day trading, the reversal is preferable. This allows the chart to reflect the reversal sooner and bring the stop-loss closer—both will serve to reduce risk. Be sure to continue the chart from day to day rather than begin anew each day. This will provide continuity. Point-and-figure is also subject to directional trading slippage.

Moving Averages

The intrinsic dependency on time, as well as its built-in lag, makes a moving average less adaptable to day trading. Traditionally, a moving average is recalculated at fixed intervals. Two "philosophic" questions must be answered with regard to intraday moving averages:

1. Do trends exist in this short time interval?
2. Does it make sense to apply a moving average to intraday prices unevenly spaced with respect to time?

Throughout this book, the conflict between noise and trends is discussed. Short time bars (for example, 15 minutes) have relatively more noise than larger bars (daily). Larger relative noise requires more time to identify a trend and a larger price move to profit from it; therefore, it may not be possible to enter a trend trade and still have enough profit potential before the end of the trading session. We will see that the span between the open and close is only about 50% of the trading range. Then it would be unlikely to capture more than 25% of the daily range after the trend has been identified, making the profit opportunity too small. Then trend trading using intraday bars will need to span a number of days to be profitable.

To be fair, there are analysts who believe a moving average is a fine tool for day trading. It has the same performance profile in the very short term as it does when applied to daily or weekly data, that is, it cuts losses short and lets profits run. However, no single calculation period seems to satisfy all markets because the difference between the slow moving Eurodollar rates and the explosive NASDAQ-100 require longer and shorter periods, respectively. Those planning to explore the use of a moving average should consider two issues:

1. Does the need to use different calculation periods constitute overfitting, or is it simply the necessary reflection of the individual market personalities?

2. Are the higher relative costs of day trading a serious obstacle to profitability, differentiating it from longer-term trend following for which transaction costs have little effect?

Choosing the Intraday Bar Size

It is always easier to use features normally provided by your quote equipment. And we normally think of intraday bars in 5-, 10-, or 15-minute intervals, starting at the top of the hour. Developing a successful trading strategy is difficult enough, but using a nonstandard bar may be adding unnecessary complications. Is it best to trade the way the great majority are trading, or is there an advantage to moving off the normal path?

Selecting the Bar Interval

While there is no *correct* bar length, traders often choose 5-, 10-, 15-, 20-, 30-minute, and 1-hour periods to chart. Some system development platforms require that the interval be evenly divided into 1 hour; therefore, a 6½-minute interval is not allowed. When using longer bars, such as 1 hour, it is important to consider the length of the last bar of the day. In trading platforms or charting software that requires that all bars begin at the start of the hour, a 1-hour bar will post its first price at the end of the first actual hour, rather than one hour after the open; the last bar will be the interval from the last whole hour to the close. In the case of the energy markets, which close at 2:30 P.M. in New York, this will make the last hourly bar 30 minutes; for the S&P pit session, closing at 4:15 P.M., and grains closing at 1:15 P.M. in Chicago, it would be 15 minutes. For a 30-minute bar, a market closing at 2:30 may have only 1 minute in the last bar. When bars are uneven in this way, the volatility and volume for each bar can no longer be compared.

There is a difference in opinion between analysts over the choice of intraday intervals. One group prefers to pick a standard number because it conforms to the way others trade; for example, you can expect orders to flow more actively on the hour. This approach may be particularly valuable if you plan to take the opposite position. Another group of traders are constantly seeking their own time interval, avoiding uneven order flow and price distortion caused by most other traders. If you are part of this group, you may choose your interval by dividing the total number of minutes in the trading session by a value that gives you equal-length bars. You may also want to find a bar length, such as 21 minutes, that represents a Fibonacci number and comes close to dividing the day equally.

Uneven Price Bars Based on Tick Count

With any fixed-length bar, there are significant variations in volume and volatility. During mid-morning or mid-afternoon, volume can be low. While normal trading in heating oil occurs every few seconds, at 12:30 P.M. there might be gaps of 10 seconds or more when there are no trades. For older data, these gaps could have lasted minutes. This creates a very large variation in volume and volatility. To avoid that, some traders will choose bars representing an *equal number of ticks* (or price changes), normally 10, 50, or 100 price

prints. Conceptually, this is similar to Richard Arms' method, *Equivolume*, where bars that are charted vary in width according to daily volume.

Uneven price bars have their own difficulties. As with equal time bars, the last bar of the day will not have the same number of ticks and will vary from day to day. Consider two examples. First, prices are equally spaced, and second, the price of a less liquid market changes every 15 seconds during more active times, then quiets to 5-minute intervals; at one time, there is a lull of 15 minutes followed by a jump in price. Is there a difference in the way we would interpret the following two patterns? Is there a difference in the way a moving average would calculate the trend?

In the table below, Pattern *A* is based on equal 1-minute bars, while pattern *B* has an equal number of ticks in each bar. Pattern *A* shows a gradual upwards movement while pattern *B* shows slow rallies and fast drops. When interpreted by a trader, pattern *B* is often considered *toppy*; it is a market struggling to go higher with sellers waiting anxiously for their opportunity. Moving averages of pattern *A* taken as shown and pattern *B* interpolated to equal intervals will have the same trend direction at each point. There is not enough information to know which is best, but it is clear that the patterns can be interpreted differently.

Pattern A: 1-minute bars	End of bar	Pattern B: Equal number of ticks
Time	**Price**	**Time**
10:00	5005	10:00
10:01	5010	10:05
10:02	5005	10:06
10:03	5015	10:16
10:04	5025	10:26
10:05	5020	10:28
10:06	5015	10:29
10:07	5025	10:44

One conceptual advantage of the equal-tick bar is that it deemphasizes the quiet periods of the day. By combining 50 price prints into one bar it may be possible to avoid trading during the middle of the day, when activity is low and many false signals occur due to a single trade of insignificant size. A fixed-interval moving average will give equal importance to these periods.

Time-of-Day Patterns

Time-of-day and *short-term patterns* are combined in the most popular forms of day trading; these types of patterns have already been discussed in detail in Chapter 15. Patterns have changed over the years because traders seek out patterns to exploit. Many of them are not as extreme, possibly because new orders, such as VWAP (volume-weighted average price), have distributed buying and selling across the day. However, there are still opportunities, and many of these center around prices breaking through key levels.

Trading Key Levels

Traders are clear about specific price levels that they believe trigger changes of direction or continued moves. They also want those moves to happen when there is sufficient activity. These *key levels* are the ones most easy to remember—yesterday's high, low, and close, today's open, high, and low, and the most recent significant high and low. Traders remember these levels because it is their job to do that. Floor traders and market makers limit their focus to a few markets, so remembering those levels is easy.

There are general patterns that are expected throughout the day, even when they do not show in the time-of-day analysis in Chapter 15. Following the opening session each day, volume drops toward midday. Along with that, volatility drops. Around midday, the high or low of today's trading range is tested. It is common for most day traders to buy the bottom of the range and sell the top. A break of either support or resistance after midday is considered a major directional change; traders quickly shift to the direction of the breakout with the expectation of holding that position for the balance of the day. This concept forms the basis for the famous *opening range breakout system* discussed later in this chapter. A classic analysis of how key levels are traded is shown in Figure 16.1. This chart, created by Walt Bressert in the 1970s, uses the important recent highs and lows and is very specific about relationships between the developing range, the previous day's range, and the time of day.

To clarify the way Figure 16.1 is to be read, we use line 1 as an example. Note that columns *A–D* are days where the open was higher than the previous close, and columns *E–H* are days where the open was lower than yesterday's close. The notation in the box alongside number 1 shows the preconditions of that row, explained at the upper right of the chart; therefore, $H_E < H_Y$ says that *the high of the entry is lower than yesterday's high.* This row shows the beginning of an *inside day.* Columns *A–D* give the trading signal as the day progresses. In the first box, 1*A*, the absence of broken lines marked *B* and *S* for buy and sell, indicate no trade occurs with this pattern for the first 30 minutes of the day.

To avoid confusion with the notation, you should replace the *E* with a *T*. Rather than saying "the high of the entry" it is clearer to say "the high of today." The use of "entry" means the "entry day," which is also today.

The second box in row 1, 1*B*, shows that a buy and sell order can be placed 30 minutes after the open. The new long is entered on an upwards break of the previous day's highs, and a new short sale is entered on a drop through the previous day's low. By midday, box 1*C*, the buy and sell signals become closer by raising the short sale to a break of the current day's low, rather than the previous day's low. In box 1*D*, 35 minutes before the close, if prices are between the current open and the current high, the buy signal is lowered to the current high and the short sale to the opening price of the day. Bressert's chart is an extremely good example of the way a floor trader reacts to changing patterns in combination with the constantly moving clock.

Trading in the Wrong Time Zone

Exchanges that are open 24 hours allow us to trade reactions in U.S. markets when, for example, an event occurs in Asia. This corrects a problem best seen in the CME's Nikkei

FIGURE 16.1 Intraday timing of market movement.

Source: With permission of Walter Bressert, Inc. http://walterbressert.com.

750

225 contract, which has a pit session from 7:00 A.M. to 3:00 P.M. Chicago time. While that allows it to trade alongside other currency contracts, the time change from Chicago to Tokyo means that the Japanese are asleep. Then any price changes in the USDJPY reflect how we see the Japanese reacting to our economic news, not how the Japanese actually respond.

Table 16.2 compares the Nikkei 225 traded on the CME in Chicago with the same contract traded on SIMEX in Singapore. During the 1990s, the overnight gap on the CME contract was larger than the daily trading range, indicating that prices jumped on the open to adjust to trading during the previous Japanese business day. More recently, the CME has included the night hours in their data, causing the two markets to appear the same. However, to take advantage of trading during the Japanese business day from Chicago, you would need to trade from 6:30 P.M. to 1:30 A.M. because the time change is 11 hours.

The S&P, also shown in Table 16.2, compares the pit session for the large contract to the popular *e*mini contract. In this case the big S&P contract has kept the same hours since 1983 while the *e*mini contract trades from 6 P.M. until the next day at 4:15 P.M., New York time. The gap represents the price move during the 1 hour, 45 minutes when the market is closed.

Table 16.3 gives a broader view of overnight gaps. Overall, it shows that gaps are getting smaller. During this period, most of the markets show significant increases in the

	TABLE 16.2		Overnight Gaps as a Percentage of the Daily Trading Range for Nikkei (CME and SIMEX) and S&P (pit session and Globex *e*mini)	
	Nikkei		**S&P**	
Year	**CME**	**SIMEX**	**Pit**	**eMini**
1994	98.0	36.3	20.1	19.9
1995	144.2	26.6	21.7	14.4
1996	123.9	25.1	24.7	4.3
1997	129.5	26.0	23.4	3.9
1998	109.6	27.6	27.1	2.7
1999	101.4	28.8	25.9	3.0
2000	86.6	36.1	24.9	4.1
2001	84.7	33.9	27.3	5.4
2002	81.5	35.9	27.6	3.9
2003	84.2	39.4	25.8	4.4
2004	57.0	42.9	25.6	3.2
2005	47.9	10.7	21.0	3.4
2006	33.6	4.4	23.2	2.8
2007	21.9	13.8	31.4	2.6
2008	21.9	7.9	33.2	3.6
2009	29.6	9.1	37.2	3.4
2010	13.5	10.7	36.9	4.1
2011	8.2	9.0	40.5	2.7

752

TABLE 16.3 Gap Openings as a Percentage of the Trading Range for Selected Futures and Stocks, 1994–2011

| | Futures Markets | | | | | | | | | | | Stock Markets | | |
| | Index | | | Rates | | Currency | | Energy | Ags | | | | | |
Year	SP	NQ	CAC	Notes	Sh Stlg	EUR	JPY	Brent	Cattle	Sugar	BAC	GOOG	GE	ABX
1994	19.9		32.5	8.5	24.7	9.4	9.5	21.5	22.4	22.0	24.2		23.0	23.9
1995	14.4		31.5	7.3	25.7	9.5	7.6	22.2	19.6	22.8	19.8		19.3	25.5
1996	4.3	27.5	33.2	7.2	22.9	9.4	8.5	22.2	21.8	22.0	25.5		21.9	26.8
1997	3.9	24.5	43.5	9.1	23.7	8.4	8.7	24.1	19.0	20.4	25.7		26.3	30.8
1998	2.7	17.8	19.8	6.8	23.8	8.1	8.1	26.9	17.8	27.7	25.6		26.2	32.5
1999	3.0	6.1	26.6	4.6	21.5	13.1	8.6	25.4	21.4	24.7	26.3		22.2	31.5
2000	4.1	3.8	24.1	7.1	20.3	7.2	7.2	23.6	18.2	26.8	20.6		23.4	26.2
2001	5.4	5.6	26.3	11.7	24.2	9.5	10.1	20.4	22.2	23.1	24.7		23.7	24.2
2002	3.9	5.4	25.6	12.9	24.7	10.4	10.3	27.4	23.5	23.6	23.9		30.1	27.4
2003	4.4	5.4	23.6	13.1	22.2	11.4	11.6	20.0	33.6	27.2	22.2		27.8	24.5
2004	3.2	4.6	28.6	9.5	18.6	9.4	8.7	16.6	26.1	25.6	25.6	1.7	26.4	31.8
2005	3.4	4.5	29.8	9.4	17.2	8.5	9.8	15.5	26.7	27.8	24.1	1.9	24.9	30.4
2006	2.8	3.2	26.6	6.8	17.2	9.1	10.4	12.1	21.7	35.2	24.8	2.7	23.3	29.8
2007	2.6	3.0	32.0	12.4	19.3	9.6	12.6	10.4	13.1	37.6	26.8	3.1	24.9	31.5
2008	3.6	4.2	31.1	13.8	19.4	14.2	16.8	11.3	14.7	14.7	31.5	5.7	32.2	30.8
2009	3.4	3.8	22.0	10.3	16.3	10.7	8.8	11.7	14.3	11.2	37.0	2.8	35.2	35.4
2010	4.1	4.4	24.6	10.3	18.0	11.3	10.3	10.2	12.1	10.2	29.4	3.5	30.6	34.2
2011	2.7	3.4	22.4	9.6	17.7	7.0	10.3	8.5	12.8	7.8	32.1	3.3	38.0	32.4
Price	**SP**	**NQ**	**CAC**	**Notes**	**Sh Stlg**	**EUR**	**JPY**	**Brent**	**Cattle**	**Sugar**	**BAC**	**GOOG**	**GE**	**ABX**
First	846	1562	2222	47.18	86.83	1.366	1.514	29.94	143.30	12.47	12.74	179.93	8.11	23.42
2011	1303.1	2317	3935	119.06	98.743	1.386	1.213	109.36	114.22	28.10	12.82	563.83	19.67	48.65

underlying price (given at the bottom of the table). Higher prices mean greater volatility, which would be seen in the trading range, but not as much of an increase in the overnight gap. Markets where the gap has clearly declined are the S&P, NASDAQ, Brent, Cattle, and sugar. The Europeans, the CAC and Short Sterling, have maintained at relatively the same level, as have the currencies. The stock market shows remarkable consistency.

What does it mean? The idea of a gap opening as a significant trading signal seems less likely in today's futures markets. Because stock market earnings as well as company press releases occur after market hours, the reaction is seen on the next open, and gaps may still be a good trading signal. This does not eliminate the opportunities caused by price shocks, which may occur during the trading day following an economic report. It does mean that there are few shocks that occur when the market is not trading.

TRADING USING PRICE PATTERNS

A trade that lasts from 1 to 3 days can be improved if short-term patterns or cycles can be found, such as those in Figure 16.1. For example, in a trending market there are outstanding weekly and weekend patterns. It is common to find that Monday's move is a continuation of that trend. By Tuesday (or sometimes Wednesday), the strength of new buyers or sellers has faded, and the market reverses due to lack of activity and some profit taking. It often stays in this state through Wednesday or early Thursday when it again resumes the trend. Friday afternoon may see a minor reversal of trend as some traders take profits and reduce risk for the weekend. In a sideways market, the Friday and Monday directions differ from the direction during the prior week and often differ from each other.

Variations on Support and Resistance

Support and resistance levels are most important to the short-term trader. If prices start to move up, slow down, and finally reverse, it is natural to consider the highest price as a resistance point. Prices are thought to have been stretched to their extreme at that level. Any subsequent attempt to approach the previous high price will be met with professional selling in anticipation of prices stopping again at the same point or slightly lower. In addition, it is common for the traders to place orders above the previous high prices and enter new long positions or closeout shorts in the event of an upwards breakout through resistance.

This method, very popular among floor traders and active speculators, tends to create and then reinforce the support and resistance price levels until they define a clear *trading range*. Within a 1- to 3-day period, these ranges can be narrow and yet effectively contain price movement. During the life of the trading range, it will continue to narrow as the levels become clear to more traders and the anticipation of a reversal at those levels becomes imminent. This is similar to an ascending or descending triangle, where the price pattern continues to narrow; only in this case the time frame is much shorter than that of traditional charting.

There are many combinations of the day's open, high, and low, along with yesterday's high, low, and close, that can be used to create a trading strategy. A good summary was shown in Figure 16.1. We can also create a strategy that uses the idea of support or resistance near the middle of the trading day, where volume and volatility often drop to low points for the day and prices have a good chance of reversing direction. If they continue to make new lows or new highs after midday, then we expect prices to continue in that direction for the rest of the day. The rules for this strategy, which we will call *TSM Midday Support and Resistance* (available on the Companion Website), using 15 or 30 minute bars, are:

1. Define "afternoon" beginning at 12:00 P.M. local time.
2. Record the high and low from the open until 12:00 P.M.
3. Record the highest close and lowest close from the open until 12:00 P.M.
4. If we are in the afternoon, then
    a. If there is no position or we are short, and the closing price of the bar is in the lower 25% of the morning high-low range, then cover shorts and go long.
    b. If there is no position or we are long, and the closing price of the bar is in the highest 25% of the morning high-low range, then exit longs and go short.
    c. If we are long and prices break below the low of the morning range, then exit longs and go short.
    d. If we are short and prices make a new high above the morning range, then exit shorts and go long.

This strategy takes advantage of both support and resistance as well as an intraday breakout, limited to the afternoons. Trades can be held into the next day, which can greatly increase the profits per trade. Figure 16.2 shows trading signals generated for the S&P futures during April and May 2011. Trades marked as *MDlong* and *MDshort* are breakouts during the afternoon, and trades marked *MDsup* and *MDres* are reversals based on failing to break support and resistance levels. In its basic form, using 12:00 P.M. as midday and a 25% trigger zone for support and resistance decisions, this strategy showed gains of $268 per trade (1 contract each, no costs) on 8,404 trades from January 2000 through mid-May 2011. There was no attempt to improve results by changing either the midday time or trigger zone, nor were any other rules tried other than those shown above. The purpose was to show that there are certain principles understood by traders that can be turned into a profitable strategy. Having said that, the S&P is a likely choice to be successful because it has more noise and volatility than other markets, contributing to its ability to reverse at a support or resistance level. Other more trending markets may do better without the midday reversal rule but with the afternoon breakout. Similarly, the opening range breakout capitalizes on a more popular approach and is discussed later in this chapter. Adding a trend filter would also improve results of interest rate and FX markets.
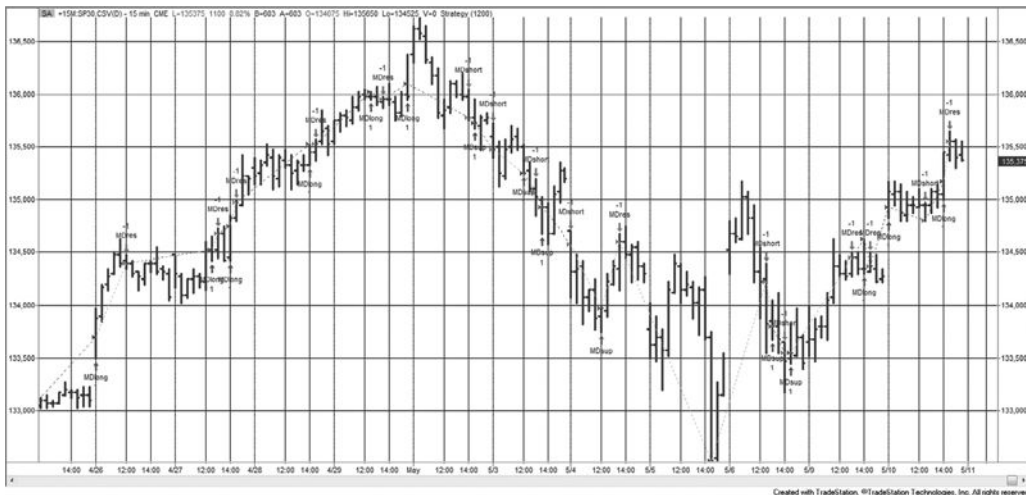
FIGURE 16.2 Sample of S&P trades using the *TSM Midday Support and Resistance* strategy, April and May 2011.

The Taylor Trading Technique

In 1950, George Douglass Taylor published his *book method* of day trading which he had been using for many years in both the stock and grain markets.[2] While this preceded the availability of a computer, the method was fully systematic and clearly written in his book. The system is intrinsically cyclic, anticipating 3-day movements in the grains. These 3 days can vary in pattern when they are within an uptrend or downtrend. Taylor's explanation of his method is thorough and includes many valuable thoughts for traders interested in working with the market full time. The summary and analysis presented here cannot replace a reading of the original material. It should also be interesting that a 3-day cycle continues to be a successful pattern for short-term trading and that this method, and variations, have had a resurgence of popularity in recent years.

Taylor developed his approach to trading through a belief that there is a basic rhythm in the market. The dominant pattern is seen to be a 3-day repetition with occasional, although regular, intervals of 4- to 5-day patterns. Taylor's cycles are based on continuous trading days without regard to weekends and holidays. The 3-day cycle varies slightly if prices are in an uptrend or downtrend. The *uptrend* is defined as having higher tops and bottoms over some selected time period such as a week, month, or season. A *downtrend* is the reverse. During an uptrend the following sequence can be expected:

1. A *buying day objective*, where prices stop declining and a purchase can be made before a rally begins.

---

[2] *The Taylor Trading Technique*, by George Douglass Taylor (1950), has been reprinted by Traders Press, P.O. Box 6206, Greenville, SC, 29606 in 1994.

2. A *selling day objective*, at which the long position is closed out.

3. A *short sale day objective*, where prices meet resistance and can be sold prior to a reversal.

Following the third day, after a short position is entered, the cycle begins again with a *buying day objective.* Because an uptrend has been identified, Taylor has given extra latitude to the long position with part of the day between the liquidation of the long (2), and entering a new short (3) reserved to allow the upwards momentum to exhaust itself. Downtrends are the opposite, expecting some added time for the downwards move to finish before the rally begins.

The actual *objectives* are extremely short-term support and resistance levels, usually only the prior day's high and low prices or occasionally the high and low of the 3-day cycle, which may be the same prices. On a buying day, the objective is a test or penetration of the prior day's low price, but only if it occurs first, before a test of the prior highs. Taylor's method is then a short-term mean-reverting technique, which looks for prices to reverse direction continuously. His belief was that speculation caused these erratic, sometimes large, cyclic variations above and below the long-term trend.

Taylor placed great emphasis on the order of occurrence of the high and low on each day. In order to buy, the low must occur first. If this is a buying day in an uptrend, a long position is entered after any lower opening, whether or not the prior low is reached. Taylor reasoned that because an uptrend is generally stronger toward the close, the first opportunity must be taken on a buying day. However, if a high occurs first and prices then decline, nearing the lows toward the end of the trading day, no long position is entered. This pattern indicates a lower open the following day that will provide a better opportunity to buy. During a downtrend, the violations of the lows, or penetrations toward the close, are more common. By waiting until the next day to purchase, the 3-day cycle is shifted to favor the short sale.

Consider the same problem with regard to closing out a long and entering a new short during an uptrend. If on the same day that the long was entered prices rallied sharply, touching or penetrating prior highs, a higher opening would be expected the next day, at which time the long position would be closed out. Because of the uptrend, a small setback and another test of the highs might be expected. It would then require another day to ensure that the strength was exhausted. If the high of the *selling day* was tested on the open of the *short sale day*, a new short would be entered immediately. If the short sale day opened lower and finished higher, no position would be taken.

Shorter reactions to price moves are expected when a long position is entered on a buy day, or on the next open, during a downtrend. If prices rally sharply on the same day, the position is closed out at a profit. This is important to remember because trading against the trend does not offer the latitude of waiting for the best moment; time is working against the position.

Reading Taylor's Book
Taylor called this technique his *book method* because he recorded all the information necessary for trading in a small 3″×5″ spiral notebook that he carried. The organization

of the book is shown in Table 16.4, an example that uses the November 1975 soybean contract. Of course in Taylor's book, there would be only one month per page due to the limitations of the page size. The first 5 columns contain the date and day (together), followed by the open, high, low, and closing prices. The first 10 days are used to determine where the cycle begins. Scanning the daily lows, circle the lowest of the first 10 days, in this case March 19. Then work backwards and forwards, circling every third low price. These are buying days. The sequences of 2 in-between days are circled in the *High* column and indicate the selling day and the short sale day, respectively. This example is especially simple because it assumes a consistent uptrend and shows no variation from the 3-day cycle.

TABLE 16.4 Taylor's Book, November 1975 Soybeans

1975		Open	High	Low	Close	D	R	BH	BU	Net
					March					
M	10	540	(x542)	535	540½					
T	11	435	560¼	(x535)	560¼	7		18¾	0	
W	12	560	(x572)	552	571½		37			
H	13	568	(x577)	559	560½					
F	14	553	559	(✓549)	549½	28		0	10	+2¾
M	17	548	(✓560)	548	548		11			
T	18	555	(x558)	538	538½					
W	19	530	546	(x529)	545¼	29		0	9	
H	20	543	(✓555)	540	547½		26			
F	21	547½	(✓567½)	547½	567½				−1½	
M	24	573	581	(✓561)	573	6½		13½	0	
T	25	573	(x575)	564	574		14			
W	26	577	(✓594)	574	594					
H	27	587	590	(✓579)	585	15		0	0	+31½
M	31	592	(✓603½)	589	595½	24½				
					April					
T	1	599	(x599)	584	591½					
W	2	589	591½	(✓575)	575¾	24		0	9	
H	3	570	(✓573)	565	566¾		0			
F	4	570	(✓578)	567	576¾				−12	
M	7	577	577	(✓562)	565¾	16		0	5	
T	8	566	(✓579½)	563	568¼		17½			
W	9	565	(✓574)	561½	573½					
H	10	571	575	(✓556)	561¾	18		1	5½	
F	11	560	(x564½)	558	563½		8½			−9
M	14	561	(x568)	555½	561½					
T	15	565	569½	(✓560½)	563½	7½		1½	0	
W	16	562	(x566)	553	559¼		5½			
H	17	560	(✓565)	558	563½					
F	18	563	563	(✓557)	559	8		0	1	−1

To judge the opportunities for buying and selling, the next columns, marked *D* and *R*, indicate the number of points in the *decline* from the short sale to the buying day and the number of points in the *rally* from the buying day to the selling day. In both columns, the differences are taken using the highs and lows only. These values represent the maximum number of points that could have been made in those trades, provided the highs and lows occurred in the proper order. An ✕ or a ✓ in the circle next to the high or low means that the opportunity to buy or sell occurred first (if an ✕) or last (if a ✓). In the case of the first, the trade would have been entered that day and in the other case, Taylor would have waited.

The next two columns, *BH* and *BU*, show the risk and opportunity of that day's prices to the buying objective. *BH* means a *buying day high* and is entered with the number of points that the day traded above the prior day's high (the *short sale day*), and *BU* shows the opportunity to *buy under*, by recording the number of points by which the day sold under the low of the prior day—the area to buy if the low occurred first. If neither situation occurred, zeros were entered.

A wide column on the right was used to indicate the net weekly change in direction by taking the difference between the prior Friday's closing price and the current Friday's price. This should be used to compare with the trading performance.

By observing columns *D* and *R* in Table 16.4, it can be seen that there was ample opportunity for profits on both the long positions and short sales, with only one case of a zero entry on April 3. The trades that would have been entered or liquidated can be approximated using the *BH* and *BU* columns and the ✕ and ✓ notations. For consistency, it might be assumed that the ✓ indicates that a position was taken on the next open. In either case, the results would have been good.

Taylor's daily method requires care in monitoring the market, which only a full-time trader can do. The order of the highs and lows must be watched as well as whether the new low is going to penetrate or fail to penetrate the prior lows. In addition, it would be helpful to combine this method with a good trend-identification technique as well as observe seasonal patterns to improve the choice of the overriding market direction.

Taylor's 3-Day Trading Method

For those who cannot watch the market constantly, Taylor offers a rigid *3-Day Trading Method*, which is a modification of his more comprehensive *book method*. The cycles remain the same, but the buying and selling objectives are entered into the market in advance. This method is expected to work on balance, as are other well-defined systems. What is primarily lost by this approach is the order of occurrence of the highs and lows; otherwise, the concept of the system remains intact.

It is interesting that this technique profits from *penetration* of support and resistance levels by assuming there will be no follow-through when the right patterns occur. Most other methods take the approach that trading in the direction of the breakout is the safer alternative. It is one of the few examples of such an approach and could only succeed in the short term, where noise, rather than fundamentals, dominate price behavior.

INTRADAY BREAKOUT SYSTEMS

The 1990s were the crowning years for intraday breakout systems. With the advances in computer speed, user-friendly software, and the availability of electronic data feeds at a reasonable cost, the very popular *N-day breakout method* became the equally popular *N-bar breakout method.* Entry and exit signals could occur during the trading session based on new highs and lows of the previous 10 bars, rather than a new daily high or low. Traders also discovered that the risk could be shockingly large if there were reversals during the day caused by wide price swings—new highs, followed by new lows, followed by new highs again. Trades were held overnight, and the *N*-bar calculations continued across days.

Even though the strategies discussed earlier in this chapter include intraday breakouts, they are not pure breakout systems. In the following section the basic opening range breakout will be explained, followed by variations and examples.

Opening Range Breakout

The *Opening Range Breakout,*[3] by far the most popular of the intraday breakout methods, gives the underlying entry and exit signals based on how far prices move from the opening price of the day. The trader is then buying when prices move up from the open by, for example, $\frac{8}{32}$ in bonds, or 25 basis points in cattle. If prices then declined to 8 ticks below the open, the long position would be reversed to a short. This approach uses the opening price as the *pivot point* for the day, claiming that price direction, relative to that point is important. All trades are exited on the close of the day.

It is also reasonable to use the close of the previous day as the pivot point. Most traders watch today's prices to see if they are higher or lower than the prior close, which connects today's price patterns to the previous day. In its basic form, the Opening Range Breakout treats one day as independent of all previous price movements. This seems unnecessarily restrictive. Toby Crabel, in his extensive study of intraday patterns, has found that combinations of inside days, low volatility, bull and bear hooks, and other patterns that precede the current day, all contribute to a selection process that improves trading. Table 16.5 shows Crabel's comparison of an opening range breakout with and without using a preceding inside day.[4] The improvement is significant. However, this was done in 1990.

Updating Crabel's Opening Range Breakout

It is not possible to implement the Opening Range Breakout exactly as Crabel's program works, but we can look at the most important features, the breakout size, the inside day,

---

[3]See Toby Crabel, *Day Trading with Short Term Price Patterns and Opening Range Breakout* (Greenville, SC: Traders Press, 1990).

[4]Toby Crabel, "Opening Range Breakout, Part 4," *Technical Analysis of Stocks & Commodities* (February 1989).

TABLE 16.5   Opening Range Breakout, % Profitable Trades

Market	Breakout	Any Day (%)	Inside Day (%)
Bonds	Open +16 ticks	60	76
	Open +8 ticks	55	74
	Open –8 ticks	56	62
	Open –16 ticks	56	66
S&P 500	Open +160 points	68	61
	Open +80 points	55	57
	Open –80 points	49	48
	Open –160 points	49	45
Soybeans	Open +10 cents	60	70
	Open +5 cents	56	67
	Open –5 cents	58	69
	Open –10 cents	63	76
Cattle	Open +50 points	65	55
	Open +25 points	58	55
	Open –25 points	58	60
	Open –50 points	63	73

Source: Toby Crabel.

and holding the trade overnight. A program that does that is available on the Companion Website as *TSM Opening Range Breakout*. In addition, the program allows the fixed-point breakout to be replaced by one that is based on recent market volatility, adding an adaptive feature to the strategy.

Using different intraday bars and data periods, all ending February 2011, four futures markets were tested, crude oil, gold, 30-year bonds, and the large S&P. Because it is not clear what the size of the breakout should be, although it is unlikely to be the size used in 1990, each market was tested over a range of fixed-point breakouts. While this optimization tends towards overfitting the data, a result that improves all or most of the previous results, and shows profits over a broad range of tests, can be consider robust. This is discussed fully in Chapter 21,

Using crude oil as an example, the test used 5-minute bars from March 2009 through February 2011, with breakouts varying from $0.05 to $1.00/bbl. Remembering that crude has been trading near $100/bbl with high volatility during this period, results are still excellent after a lapse of 20 years. In the leftmost panel, the basic rules are *no overnight position and all days* (no inside day restriction), and all breakout sizes showed profits (without costs). When the breakout point size is small the profits per contract (*Prof/Tr*) is too small, but as the breakout size increases, the profits per contract approach $100. (See Table 16.6.)

In the middle panel, holding the trade overnight increases the profits per contract in some cases, especially with small breakout criteria, but overall shows a lower profit factor (*ProfFact*), a measure of reward to risk. The last panel is the most interesting. Holding trades overnight and trading only after an inside day, the number of trades are greatly

TABLE 16.6 Crude Oil, 5-Minute Data, Opening Range Breakout Tests (no costs)

Each panel shows the Net PL, number of trades, percentage of profitable trades, win/loss ratio, profit per trade (contract), and profit factor (gross profits divided by gross losses).

Crude 5m Breakout	No Overnight, No Inside Day						Overnight, No Inside Day						Overnight, Inside Day					
	Net PL	Trades	Prof Tr	W/L	Prof/Tr	ProfFact	Net PL	Trades	Prof Tr	W/L	Prof/Tr	ProfFact	Net PL	Trades	Prof Tr	W/L	Prof/Tr	ProfFact
1.00	18060	221	46.15	1.66	81.72	1.48	14720	222	41.44	1.53	66.31	1.08	23320	25	44.00	1.95	932.80	1.53
0.95	22290	234	49.57	1.59	95.26	1.59	13750	238	41.18	1.52	57.77	1.07	19690	27	40.74	2.08	729.26	1.43
0.90	19720	253	49.01	1.46	77.94	1.43	22740	250	43.60	1.45	90.96	1.12	22040	27	40.74	2.04	816.30	1.49
0.85	24090	262	51.15	1.39	91.95	1.52	15830	266	43.23	1.42	59.51	1.08	20990	29	41.38	2.02	723.79	1.43
0.80	21930	284	51.41	1.31	77.22	1.40	30140	280	43.93	1.47	107.64	1.16	32160	29	48.28	1.82	1108.97	1.82
0.75	20010	304	50.66	1.26	65.82	1.32	38510	296	43.58	1.53	130.10	1.20	31960	31	48.39	1.91	1030.97	1.79
0.70	24890	323	52.63	1.17	77.06	1.36	33100	320	42.81	1.54	103.44	1.16	39680	31	51.61	1.94	1280.00	2.06
0.65	25680	346	54.62	1.09	74.22	1.34	36230	340	40.88	1.68	106.56	1.17	41080	33	48.48	2.21	1244.85	2.08
0.60	19890	382	52.88	1.07	52.07	1.22	19480	382	39.01	1.69	50.99	1.08	47160	43	46.51	2.58	1096.74	2.24
0.55	16950	423	50.35	1.13	40.07	1.16	14630	426	38.26	1.71	34.34	1.06	47340	47	42.55	2.95	1007.23	2.19
0.50	12430	468	48.50	1.14	26.56	1.10	14580	470	36.38	1.83	31.02	1.06	29060	55	36.36	2.71	528.36	1.55
0.45	13080	511	48.34	1.14	25.60	1.10	19890	512	34.96	1.99	38.85	1.07	39380	59	37.29	3.08	667.46	1.83
0.40	2310	587	45.66	1.19	3.94	1.01	-13060	602	33.22	1.93	-21.69	0.96	47480	61	37.70	3.40	778.36	2.06
0.35	6280	656	44.36	1.26	9.57	1.04	5210	662	32.63	2.09	7.87	1.02	51000	63	39.68	3.26	809.52	2.15
0.30	12390	762	43.44	1.36	16.26	1.06	1850	752	30.85	2.25	2.46	1.01	55720	73	35.62	4.04	763.29	2.24
0.25	14590	878	41.80	1.47	16.62	1.07	21390	852	28.87	2.62	25.11	1.07	61740	73	38.36	4.01	845.75	2.49
0.20	15810	1061	37.61	1.75	14.90	1.07	27450	1044	25.38	3.17	26.29	1.08	62160	89	32.58	4.96	698.43	2.40
0.15	25030	1355	32.77	2.23	18.47	1.10	28210	1326	21.87	3.84	21.27	1.08	65820	107	28.04	6.41	615.14	2.50
0.10	18160	1872	25.21	3.13	9.70	1.06	23130	1842	17.37	5.03	12.56	1.06	57160	149	22.15	7.44	383.62	2.12
0.05	14800	2945	16.98	5.10	5.03	1.05	7970	2928	11.17	8.10	2.72	1.02	51060	235	14.47	11.31	217.28	1.91

reduced but both the profits per contract and the profit factor soar. Then an inside day is clearly a strong selection requirement, but only 10% of the days satisfy this condition, which is small.

Consider the three other markets that were tested, gold, Treasury bonds, and the S&P. Table 16.7 shows only the profits per contract. For gold and bonds, holding the position overnight increased the size of the returns at about the same number of trades. However, the profits are still too small. When the inside day condition is added, profits increase significantly in all markets along with a general reduction in trades to about 10% of the original number.

If you consider that intraday trading may give at least one new signal each day, any method that helps select the better trades is welcome. The intraday trading profile of many small profits and losses allows you to be more selective without fear of missing *the* one big trade of the year, which is a problem typical of a long-term trend-following system. If you had 250 trades per year, 2,500 in 10 years for each market, you could randomly skip 50% of the trades, or stand aside at any time, and not be concerned about lost opportunity. If reduced volatility is the key to selection, there are other possibilities that might not restrict the number of trades as severely.

Raschke Trades Crabel

Linda Raschke, one of the stars of intraday trading, working with Laurence Connors, takes her own approach to Crabel's range contraction.[5] Raschke isolates the inside day concept, focusing on a 4-day interval, NR4, where the 4th day has a smaller high-low range than the preceding 3 days. It does not have to be an inside day. The trading rules are then:

1. On the day following the NR4 day, place a buy stop 1 tick above the high of the NR4 day and a sell stop 1 tick below the low of the NR4 day.

2. If a new signal occurs—for example, a buy—and prices reverse on the same day, moving through the level of the sell stop, close out the long and sell short.

3. Use a trailing stop to avoid giving back profits.

4. If not profitable in 2 days, then exit the trade on the close of the second day.

This method can be tested using *TSM Raschke Opening Range Breakout*, available on the Companion Website. The only missing feature is the trailing stop, which was not defined.

Adding Volatility

There are two ways that trade selection could be improved: by making the breakout entry criterion more flexible and by filtering the trades when the shorter-term volatility is less

---

[5]Laurence A. Connors and Linda Bradford Raschke, *Street Smarts: High Probability Short-Term Trading Strategies* (Malibu, CA: M. Gordon Publishing Group, 1995).

TABLE 16.7 Profits per Contract for Gold, U.S. 30-Year Bonds, and S&P for Three Conditions

Gold				U.S. 30-Year Bonds				S&P 500			
Breakout	No Options	Overnight	O/N + Inside	Breakout	No Options	Overnight	O/N + Inside	Breakout	No Options	Overnight	O/N + Inside
3.00	4.68	2.82	117.48	1.00	−18.29	35.89	192.11	10.0	−21.14	−40.58	−22.22
2.80	3.55	9.00	134.64	0.95	−7.33	37.75	126.02	9.5	−21.14	−40.58	−22.22
2.60	8.87	16.40	65.80	0.90	−6.53	38.76	308.65	9.0	−23.06	−28.62	12.29
2.40	9.74	18.48	92.00	0.85	−13.55	38.02	560.17	8.5	−23.06	−28.62	12.29
2.20	12.09	14.15	40.53	0.80	−1.42	38.54	450.55	8.0	−20.25	−29.64	−58.33
2.00	5.84	11.05	46.19	0.75	−11.13	40.42	404.85	7.5	−20.25	−29.64	−58.33
1.80	9.03	13.36	66.53	0.70	−9.05	40.68	332.66	7.0	−14.88	−9.97	142.09
1.60	6.55	12.92	57.06	0.65	−22.78	38.85	360.60	6.5	−14.88	−9.97	142.09
1.40	6.11	14.81	21.62	0.60	−17.86	38.07	304.75	6.0	−15.61	−14.96	143.42
1.20	3.10	13.86	11.38	0.55	−16.98	38.83	204.76	5.5	−15.61	−14.96	143.42
1.00	1.69	8.48	60.65	0.50	−6.86	38.39	112.95	5.0	−19.13	−29.62	216.67
0.80	3.78	5.04	68.07	0.45	−3.01	38.65	108.80	4.5	−19.13	−29.62	216.67
0.60	1.14	4.72	58.53	0.40	−3.75	39.26	156.54	4.0	−17.53	−20.84	144.11
0.40	1.09	2.18	25.66	0.35	−3.41	38.79	169.26	3.5	−17.53	−20.84	144.11
0.20	−0.11	2.93	39.38	0.30	4.42	38.10	133.59	3.0	−16.61	−25.72	135.86
				0.25	5.89	36.44	146.46	2.5	−16.61	−25.72	135.86
				0.20	3.28	34.64	161.35	2.0	−21.19	−25.24	45.02
				0.15	1.88	31.33	51.85	1.5	−21.19	−25.24	45.02
				0.10	0.90	28.23	52.79	1.0	−23.01	−26.03	27.05
				0.05	−3.14	20.63	21.57	0.5	−23.01	−26.03	27.05

than the longer-term volatility. A breakout that uses the same fixed value for the threshold will not work over time. The breakout value must change over time. Retesting and changing it periodically rarely works. It is better to find a way that automatically adapts to changing price levels and underlying volatility.

A volatility breakout can be created using the following rules:

1. Find the average true range of the past n bars. This test will use 20 bars, but it is best to use the number of bars in a day, a way of removing the daily cycle from the process.

2. Multiply the average true range by a factor. That result is added to or subtracted from the open to create the breakout thresholds. The following test uses a maximum factor of 2.0, but that could have been as large as 4.0.

Results are shown in Table 16.8, which can be compared to Table 16.6. There are more trades in this new test, but the general shape of returns is similar. Most important, this method will adapt to changing market conditions. The volatility breakout feature is an option in the program *TSM Opening Range Breakout*, available on the Companion Website.

If lower volatility is more desirable, then trades can be filtered using historic volatility, which is described as having more autocorrelation than price; that is, volatility has more persistence than price. To take advantage of the expansion and contraction of historic volatility and to emphasize the consistency with Crabel's method, trades would only be entered if the 6-day historic volatility is less than the 100-day historic volatility, where historic volatility over the past n days is measured as

$$HV_i = \ln\left(\frac{P_i}{P_{i-n}}\right)$$

Once the trade has qualified, all other rules are the same. Note that this will reduce the number of trades by 50%.

Breakout Ranges Based on Time

Instead of a breakout that occurs when prices move a fixed number of points away from the opening price, or a threshold based on volatility, a popular variation is to establish an opening range based on the first 60 minutes of trading.[6] Based on the high and low established after the first hour of trading, a new high or low would signal a long or short position. This is a trade-off between avoiding the erratic movements of early trading and entering as early as possible into a price move.

Combined with the *First Hour Breakout* is a calculation of the day's trading range used for reversing your position or taking profits. The projected range high is calculated as the average *daily* range added to today's low price after the first hour of trading; the

---

[6]Malcolm McNutt, "First-Hour Breakout System," *Technical Analysis of Stocks & Commodities* (July 1994).

TABLE 16.8 Crude Opening Range Breakout Using Volatility to Determine the Breakout Criteria

Crude 5m Breakout	No Overnight, No Inside Day					Overnight, No Inside Day					Overnight, Inside Day				
	Net PL	Trades	Prof Tr	Prof/ Tr	ProfFact	Net PL	Trades	Prof Tr	Prof/ Tr	ProfFact	Net PL	Trades	Prof Tr	Prof/ Tr	ProfFact
2	20840	486	49.6	43	1.18	41980	492	37.8	85	1.17	64020	59	37.3	1085	2.69
1.9	19950	518	48.6	39	1.16	35000	526	37.1	67	1.14	73110	61	37.7	1199	3.21
1.8	19520	542	48.3	36	1.15	27830	552	37.0	50	1.11	55020	65	35.4	846	2.23
1.7	18290	577	47.1	32	1.13	31050	580	36.2	54	1.12	53540	71	35.2	754	2.14
1.6	23210	604	47.4	38	1.16	42680	604	35.3	71	1.16	56210	71	35.2	792	2.24
1.5	23840	642	46.9	37	1.16	35860	636	34.4	56	1.13	57490	73	35.6	788	2.28
1.4	15910	694	44.7	23	1.09	33590	678	33.0	50	1.12	59000	75	36.0	787	2.33
1.3	13440	741	43.6	18	1.07	42350	714	32.2	59	1.15	60670	77	35.1	788	2.39
1.2	14450	790	42.9	18	1.08	37620	774	31.1	49	1.12	63170	77	35.1	820	2.50
1.1	5160	866	40.6	6	1.02	17980	848	29.0	21	1.06	61560	83	34.9	742	2.42
1	9460	922	39.5	10	1.04	22390	906	27.6	25	1.07	61960	87	33.3	712	2.41
0.9	18550	999	38.0	19	1.08	22910	984	26.0	23	1.07	62970	93	32.3	677	2.43
0.8	12990	1104	36.0	12	1.06	21190	1088	24.3	19	1.06	63780	99	30.3	644	2.44
0.7	19830	1223	34.5	16	1.08	31790	1202	23.6	26	1.09	62520	115	26.1	544	2.34
0.6	17620	1406	31.9	13	1.07	26550	1384	21.7	19	1.07	62810	133	24.8	472	2.32
0.5	18960	1636	28.1	12	1.07	20730	1606	19.2	13	1.05	65730	137	24.1	480	2.44
0.4	24660	1895	25.1	13	1.08	29090	1862	17.0	16	1.07	56460	171	19.9	330	2.09
0.3	16600	2313	21.2	7	1.05	18290	2288	14.4	8	1.04	49460	207	16.4	239	1.87
0.2	18910	2871	17.4	7	1.06	16190	2844	11.5	6	1.04	50550	253	13.4	200	1.90
0.1	6930	3747	13.6	2	1.02	−4250	3712	8.9	−1	0.99	48680	335	10.4	145	1.82

low is that same range calculation subtracted from the high of the first hour. Both a program and indicator for testing this method, *TSM 1st Hour Breakout* are available on the Companion Website. In the practical application of this method, the first-hour range must be penetrated by at least 20 points for the S&P. If the first-hour range is nearly as large as the average 10-day range, which is used for the exit target, then substitute the anticipated high and low for buy and sell signals, rather than a breakout of the first-hour range. Other risk controls are also recommended, such as a trailing stop, a fixed stop, and reversal if prices change direction and the opposite signal is given.

Raschke and Connors' Version of the First-Hour Breakout

In order to clarify the switching from a *buy day* to a *sell short day*, Raschke and Connors created their own indicator called *Momentum Pinball*.[7] The underlying indicator, called the LBR/RSI™, is a 3-day RSI of a 1-day rate of change; therefore, the rate of change (which makes price changes more sensitive) reduces the lag in the indicator. For a *buy* signal,

- On the previous day, the indicator must have a value less than 30.
- On the current day, place a buy stop above the high of the first hour's trading range.
- Once filled, place a sell stop at the low of the first hour's trading range.
- If the trade gets stopped out, reenter if prices reverse and penetrate the original buying threshold.
- If the trade is profitable at the end of the day, carry it overnight.
- If the next open moves in the profitable direction, exit on the open.

Sell signals are the opposite of buy signals, with the indicator posting a value greater than 70 to set up the trade. While this is clearly an opening range breakout, the use of a 3-day momentum ties back to Taylor. Figure 16.3 shows an example of this method applied to the *e*mini S&P during December 2010. The top panel shows the 30-minute S&P bars, the bottom panel has daily bars, and the middle is the indicator based on the daily bars.

Testing this method on the *e*mini S&P for 2000 through May 2011, the results showed a gain of $703 per contract on 1734 trades (no costs were used). Considering the method was published in 1995, that is quite good. We should expect markets that are more trending than the S&P to perform better. A program, *TSM Momentum Pinball*, is available on the Companion Website to allow more testing.

Mark Fisher's Opening Range Breakout

In *The Logical Trader*,[8] Mark Fisher, an independent trader and founder of the largest clearing firm at the New York Mercantile Exchange, presents the most in-depth description of an intraday trading method yet published. The method itself, primarily an opening

---

[7]Laurence A. Connors and Linda Bradford Raschke, *Street Smarts: High Probability Short-Term Trading Strategies* (Malibu, CA: M. Gordon Publishing Group, 1995).

[8]Mark B. Fisher, *The Logical Trader* (Hoboken, NJ: John Wiley & Sons, 2002).
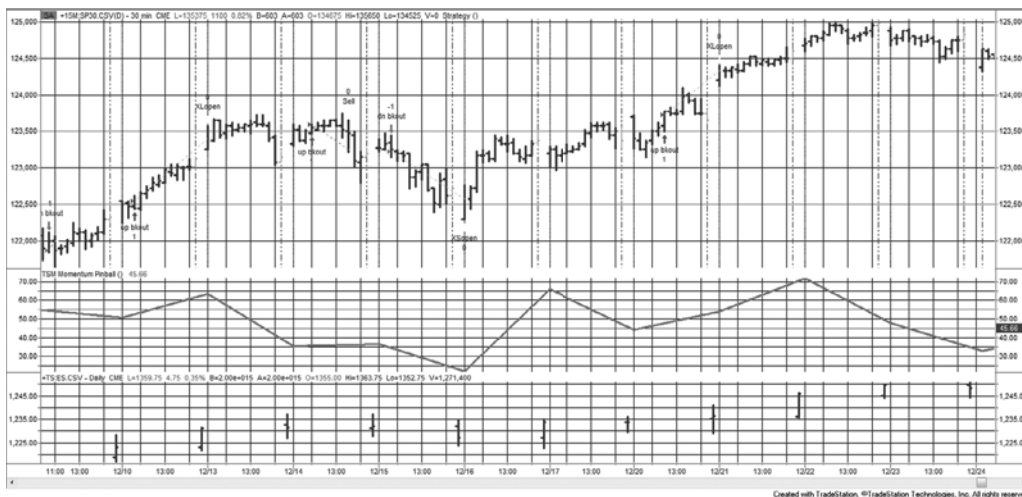
FIGURE 16.3 Using the LBR/RSI™ indicator to trade a first-hour breakout.

range breakout, is much more complex, with many additions developed over years of practical experience. And, while the techniques may be applicable to many day traders, some of the trades will target very small profits, only possible for a floor trader with very low costs and little slippage, but not practical for most screen traders. This summary only addresses the introductory concepts.

Fisher begins by defining the time interval to be used as the opening range. For stocks that interval is usually 20 minutes, and for futures somewhere between 5 and 30 minutes, depending on the market. At this point, the method is essentially an opening range breakout based on time. Because these are expected to be day trades, the time needed to recognize the opening range is kept short to allow the trade to capture as much of the daily price move as possible. It would be best to use no more than 5-minute bars for this approach. Because of the short opening range time interval, an additional requirement is that signals are confirmed by having prices remain above the breakout lines by half the time of the opening range. Therefore, for a 20-minute opening range, the price must stay above the high of that range for 10 minutes before you can buy.

We will call the high and low of the opening range *ORH* and *ORL*, and the time needed to establish the opening range as OR. Another range based on a percent of the opening range *ORH* – *ORL* (for example, 100%) is used to draw a line above *ORH* and below *ORL*, called +*A* and –*A*. Another set of lines based on a different percentage of *ORH* – *ORL* (for example, 150%) are called +*C* and –*C*; therefore +*C* and –*C* are above and below +*A* and –*A*. Using these calculations, there are five initial trading scenarios for buying (they are opposite for selling):

1. *Buy* when prices move above +*A* and remain above +*A* for a time equal to ½ × OR.
2. Close out the long trade if prices fall one tick below *ORL*.

3. *Sell short* if prices fall below $-C$ and remain below that value for $\frac{1}{2} \times OR$.

4. Once reversed to short, a stop-loss is placed one tick above *ORH*.

5. If prices do not prove profitable within a time interval equal to $1 \times OR$, then exit the trade.

Figure 16.4 shows the sequence possible beginning with a buy signal. If $+A$ is penetrated first, your trading bias is bullish; if prices reverse and fall through $-C$, then your bias becomes bearish. Although the time needed to confirm an anticipated trade is given as $\frac{1}{2} \times OR$, and the time to confirm an existing trade is $1 \times OR$, these times can be chosen by each trader. In the case of confirming a trade, the time might be 1.5 or $2.0 \times OR$.

Chosen carefully, prices will consistently find support and resistance at *ORH* and *ORL* once $+A$, $-A$, $+C$, or $-C$ have been penetrated. The selection of *OR*, the time interval,
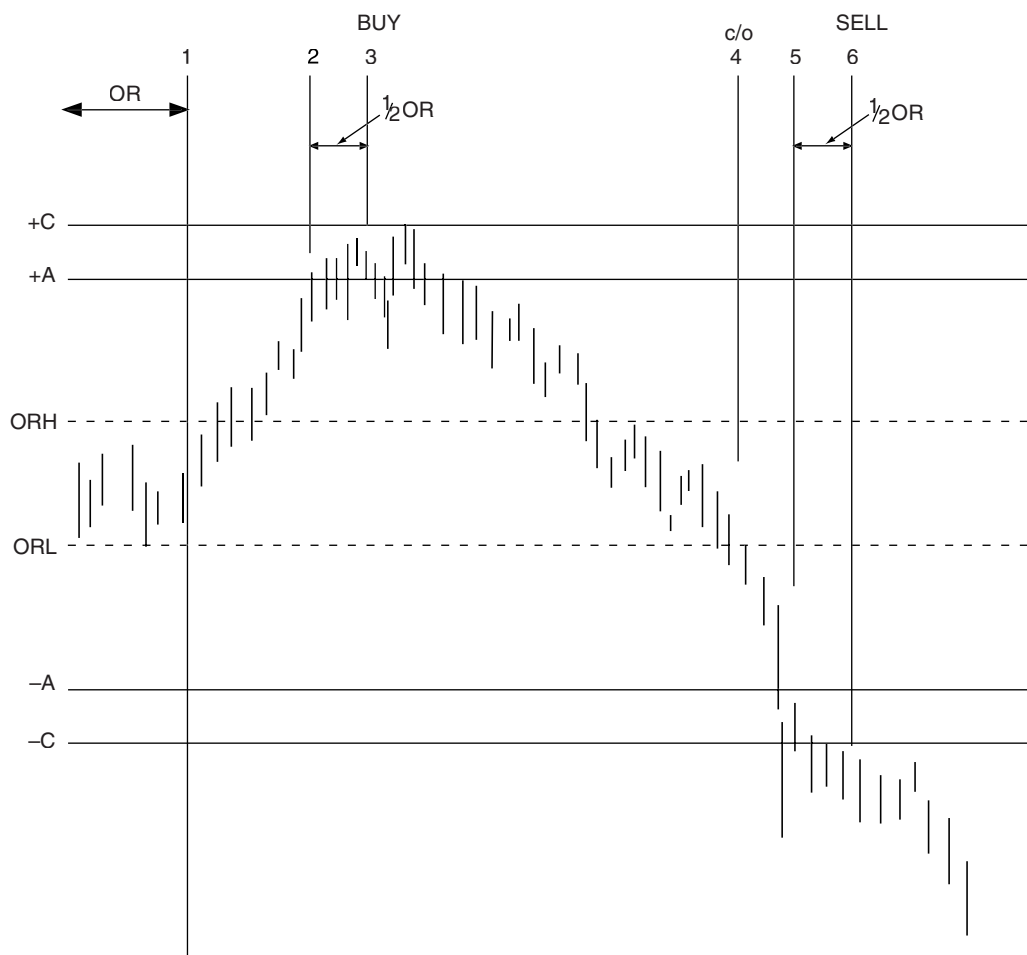
FIGURE 16.4   Fisher's buy-and-reverse scenario.

is most important because the size of the opening range is the basis for the *A*s and *C*s and the confirmation of a trade.

Profit-Taking and Resets
As with all day trading, profit-taking plays an important role. The erratic movement of intraday prices encourages profit-taking whenever there is an extended move; although not specifically discussed, profit targets can be also be set using a multiple of *ORH – ORL*. After a long position has been closed out by reaching a profit target, and prices fall back below +*A* for some time interval, a long can be reset when +*A* is broken again with a stop placed 1 tick below *ORL*.

Pivot Ranges
In order to establish a bullish or bearish bias for the day, an additional *pivot range* is calculated based on the previous day's movement.

$$Daily\ pivot\ price\ (DPP_t) = (High_{t-1} + Low_{t-1} + Close_{t-1})/3$$

$$Average\ high\text{-}low\ (AHL_t) = (High_{t-1} + Low_{t-1})/2$$

$$Daily\ price\ differential\ (DPD_t) = DPP_t - AHL_t$$

$$Daily\ pivot\ high = DPP_t + DPD_t$$

$$Daily\ pivot\ low = DPP_t - DPD_t$$

The daily pivot range can be used as a bullish or bearish bias. If the previous close was above today's pivot range high, then prices are considered bullish; if below the pivot low then it is bearish. To confirm a bullish bias, today's price movement must find support at the daily pivot low and not trade below that level. If prices break through the pivot low, they are expected to make a significant move lower; therefore, the pivot range acts as strong support and resistance. Price movement that breaks the pivot range as well as other support and resistance levels is considered more significant.

Importance of Key Ranges
Once you have found the best time to establish the opening range, then calculated the *A* and *C* levels, and finally the *pivot range*, you have a number of key levels that provide strong indication of support and resistance for the day. Penetration of those levels, and confirmation based on staying through those levels, are the primary buy and sell signals; however, these key levels are often a point of indecision. Prices can hesitate at the moment of penetration; it is also common to move through those key levels and then retreat. Fisher's time confirmation recognizes this problem.

Another entire set of rules has been written to *fade* these support and resistance levels. That is, a downward penetration of a support level (–*A*, –*C*, or pivot low) that fails is likely to be followed by a minor rally, perhaps back to the bottom of the opening range, *ORL*. These trades can provide frequent opportunities to enhance trading returns;

however, profits can be small on days with low volatility. Screen traders will need to assess for themselves whether their cost structure and execution ability allows them to take advantage of these opportunities.

Filtering the Opening Range Breakout

Larry Williams, in his day trading seminar,[9] offers improvements to the classic opening range breakout, applied to the S&P 500, using his own filters to select a better set of trades. He begins with a modification of his %R oscillator (also the 10-day raw stochastic),

$$\%R \ Index_t = \frac{Close_t - \text{Lowest}\,(Low,10)}{\text{Highest}\,(High,10) - \text{Lowest}\,(Low,10)}$$

To set up a buy signal in the S&P,

- %R Index must be below 20% and today's close must be lower than the close 7 days ago.
- Today's close must be higher than the previous close or the previous two closes to avoid buying into falling prices.
- Today must not be an inside day.

When adapted to bonds, and perhaps other interest rate markets, Williams wants more trend, therefore today's close must be higher than the close 4 days ago.

A second S&P method exploits a different set of patterns, capitalizing on the tendency to rally on Mondays and Tuesdays:

- The previous day must not be an inside day with a close greater than the low of the day before (2 days ago) or the day before that (3 days ago).
- Today's close must be less than the close 6 days ago.

For bonds, today's close must be greater than the close 15 days ago.

The actual buy signal comes when prices penetrate a band formed by adding 20% of the previous day's trading range to today's opening price. Williams also believes in the benefit of filtering trades using traditional seasonality. While they admittedly do not work all the time, they make the trader aware of the potential risk and provide a helpful edge.

Afternoon Breakouts and Midday Reversals

Early breakouts are most popular, but breakouts after midday also indicate a new price direction that should continue for the rest of the day. Because this cuts the potential profit by halving the time for the market to move, holding the trade overnight is essential.

The drop in volume and volatility at midday in most markets allows for the possibility of a reversal (if midday is also near the low or high of the day) or, if a breakout, a

---

[9] James T. Holter, "Delivering the Day-Trading Experience," *Futures* (August 2001).

continued directional move. A program *TSM Midday Support and Resistance* is available on the Companion Website to test this. This method was discussed in an earlier section "Variations on Support and Resistance."

This program can also select only the afternoon breakout signals and not the support and resistance trades. Both the midday time and the support and resistance zones can be changed.

Special Set-Up Patterns for Stocks

Both intraday patterns formed near the beginning of the day and daily price patterns that precede today's opening can be used to find the best intraday trading opportunities.[10] Rudd selects the highest-probability buy signals, always using 5-minute data. The patterns he describes are:

1. *Dip and rally.* Prices dip shortly after opening, lasting less than 1 hour, then rally back to reach the highs of the first few bars. They move sideways for some time, testing the highs without making new highs. Place an order to *buy* the breakout. The longer prices remain just under the high of the day, the more reliable the signal.

2. *Consolidation.* Prices move sideways for about 1 hour after establishing an opening range during the first few bars. Highs during the consolidation period tend to hug the highs of the day. *Buy* a breakout of the range. A narrow range is expected to give a more reliable signal.

3. *Delayed consolidation.* Prices open lower, then move higher during the first hour, settling at a resistance level. Prices test the highs but do not penetrate. Place an order to *buy* a breakout after resistance has held for at least half an hour. Longer consolidations are better.

4. *Early breakout.* After a 15-minute narrow opening trading range, *buy* a breakout on the next bar. Prices may show a series of shorter consolidations and smaller breakouts.

5. *Reversing after a false breakout.* If, after buying a breakout of a narrow consolidation pattern (2), prices fall back into the original trading range, *close out* the long and sell a breakdown of support.

Day Trades Following a Wide-Ranging Bar

Days in which the ending pattern is a *wide-ranging bar* (a high-volatility outside bar) that closes near the high of the day offer two special buying opportunities for day trading (selling is simply the reverse pattern) on the following day.

1. *Breakout of a higher open.* If, during the first 15 minutes of the day, prices open above the high of the previous wide-ranging bar, trade lower to test the high of

---

[10]Barry Rudd, *Stock Patterns for Day Trading and Swing Trading* (Greenville, SC: Traders Press, 1998).

the wide-ranging bar, then begin to rally, buy the breakout of the current day high. Greater reliability should be expected if there is a well-defined upwards trend in daily prices.

2. *Selling a gap open.* A substantial gap open should be sold if prices begin to fade during the first 5 minutes of trading.

Reversal Bar Set-Up

Rudd also looks for a reversal set-up in two preceding days, in order to take advantage of the completion of the actual reversal on the third day. The set-up must involve the following three events:

1. The *first daily bar* must close near its lows.
2. The high of the *second daily bar* must not be much higher than the low of the first daily bar, and must close near its high.
3. The high of the first daily bar should be at least 1 point above the high of the second daily bar, because the difference represents the profit potential for the next day trade.

Following a successful set-up, the day trade is entered on Day 3. The entry signals should be one of the original set-up patterns, such as a *dip and rally* or *consolidation.* A large range in Day 1 provides better opportunity.

Fading the Opening Range

A twist on the opening range breakout is based on the idea that the open and close of the day are the most likely times for the daily high and low. If you combine that with some minimum opening gap, then it is possible to fade the opening direction. If that does not turn out to be one of the extremes of the day, the trade is reversed when a new high or low occurs.

Fading the market (going against the current direction of prices) is more likely to be successful if it targets a market with high noise, that is, one that has a tendency to reverse often. That makes the S&P the best candidate and the interest rates the worst. This strategy, called *TSM Intraday High-Low*, is available on the Companion Website. The large S&P contract was used because it has the old pit session hours, which exaggerates the gap. The rules are:

1. Define a minimum gap (in point or percentage returns). For the S&P, a gap of 0.50% is used.
2. Define the bar on which a signal will be taken if the opening gap is satisfied.
3. *Buy* on the designated bar if the gap is satisfied and the gap direction is down. Set a stop and reverse as the low of the day at the time of entry.
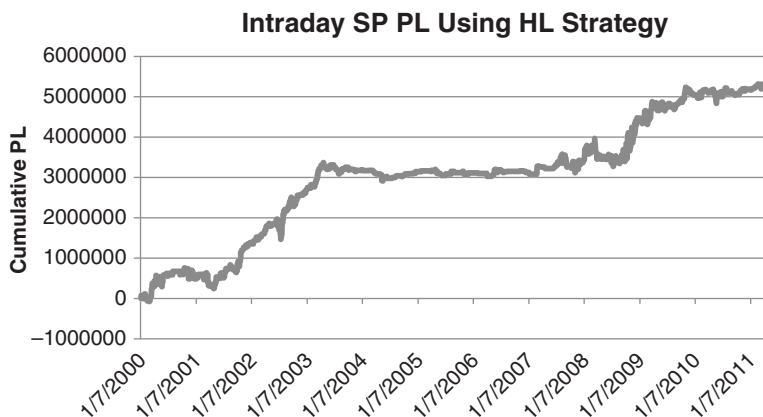
FIGURE 16.5 Cumulative PL of the S&P Intraday High-Low Strategy.

4. *Sell short* on the designated bar if the gap is satisfied and the gap direction is up. Set a stop and reverse as the high of the day at the time of entry.

5. *Exit* on the close of the day.

Using 30-minute bars from January 2000 through May 2011, a test of the S&P showed more than 1600 trades with an average profit of $32 per contract, without costs (cumulative profits shown in Figure 16.5). While it does show that this is a viable strategy, trades would need to be filtered for volatility to increase the size of the returns. This could be done with a relative measure, such as short-term divided by long-term volatility, or even an inside or low volatility day preceding the gap.

Short-Term Gap Patterns

A traditional gap is where the market opens above yesterday's high and does not trade back down to that high. Both Tom DeMark and Larry Williams have defined a *lapse* as a less restrictive case where the market opens above the previous close and remains above that price.[11]

In Williams' method, based on a *true upward gap*, you enter a new short position on a stop when the market pulls back to the previous high, or buy when a lower gap open rallies to the previous day's low. The trade is exited at the close of the day.

A more extensive program, *TSM Intraday Gaps*, can be found on the Companion Website. Entry choices for an upwards gap are

1. *Buy* on the close of the gap opening bar.

2. *Sell* short if prices pull back to the high of the previous day.

3. *Sell* short if prices pull back to the close of the previous day.

---

[11] Murray A. Ruggiero, "The Science of Pattern-Based Trading," *Futures* 2004.

Exit choices are

1. *Exit* on the close of the day.
2. *Exit* on the next open.

Using intraday data of 15 to 30 minutes, it seems that entry option (3) and exit option (2) are best. A penetration of the close is more important and exiting on the next open adds risk but tends to be rewarded.

INTRADAY VOLUME PATTERNS

Volume can be interpreted as potential volatility or lack of interest, corresponding to whether it is high or low. When volume is higher than normal, then traders are active, and something is happening in the market. Increasing volume during the day could be a sign to set a new position, but relative volume can be misleading. As we saw in Chapter 12, there is a *U*-shaped pattern of volume during the day. Traders are less active at midday; therefore, volume systematically declines toward midday, then increases toward the close. If a strategy requires increased volume (or even increased volatility), then it will always be buying in the afternoon and rarely before midday. Also, with low volume at midday, those signals would be unreliable.

Figure 16.6 shows the intraday tick volume patterns in the bottom panel for 5-minute crude oil data. The solid line is the average of volume *for that bar each day*. The histogram is the actual tick volume for that day. While this illustration only shows seven days of trading, the idea is that higher-than-average volume at the beginning of the day sets the pattern for the day, generally one with more volatility. With lower volume, we could
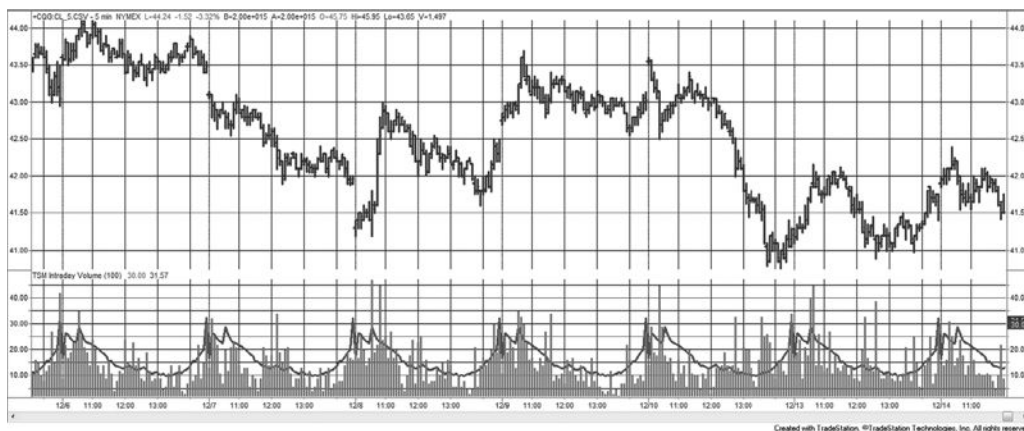
FIGURE 16.6 Intraday volume patterns for crude oil, December 2004. Today's 5-minute tick volume compared to the average for that bar.

expect a reversal, or simply a day with no fundamentally based movement, often resulting in a slow drop in price. This indicator is available as *TSM Intraday Volume* on the Companion Website. Be sure to set the *style* for the average as a *histogram* and the *style* for the current volume as a *line*.

INTRADAY PRICE SHOCKS

A price shock is a high-risk, unpredictable event. The shock can be in either direction, but always seems to be in a way that loses money for you. Although no one cannot reasonably expect to profit from anticipating the shock, there are strategies that try to take advantage of the reaction to the shock. Shocks in daily price movements were discussed in Chapter 14. Here, intraday shocks will be the focus.

Price shocks can occur in any market, but some are more interesting than others. Interest rates have frequent small jumps when economic reports are released, and a few large jumps when the Fed has surprised the market by raising or lowering rates in a unscheduled meeting. Crude oil, however, has been at the center of geopolitical events in the Middle East, and has the occasional large jump that offers unique trading opportunities.

A price shock will be defined as an intraday bar that is at least 3 times the size of the previous average bars. Of course, for the previous study of volume we should recognize that measuring the average is a problem. Is it the average of the last n bars, the average of the past (exactly) 3 days of bars, or the average of this bar over the past n days? When the volume is very low near midday, should the criteria be more than 3 times the size? Would volume work as well as volatility? If we look at Figure 16.6, it appears that volume that is twice the average of the same time bar isolates an intraday shock.

By changing the intraday volume program in the previous section to show intraday volatility, we can compare it with the intraday volume results and get a visual confirmation of whether it isolates the types of price moves that we can trade. In the bottom panel of Figure 16.7, based on 20-minute gold futures, there are a number of spikes in the daily volatility compared to the histogram showing the 100-day (average off) volatility. Note that there is a large spike each day on the open and many of the largest spikes of the other days are also on the open. This is also true of the stock market, where many spikes occur on the open, and the remaining in reaction to economic reports and news released during market hours. Intraday volatility can be plotted using the program *TSM Intraday Volatility*, available on the Companion Website.

Because we cannot predict the jump, any profits will be earned by trading the reaction to large intraday moves. We can create a strategy with the rules:

1. Calculate the average true range for each bar.
2. Calculate the average-off true range for each bar, based on a 100-day period.
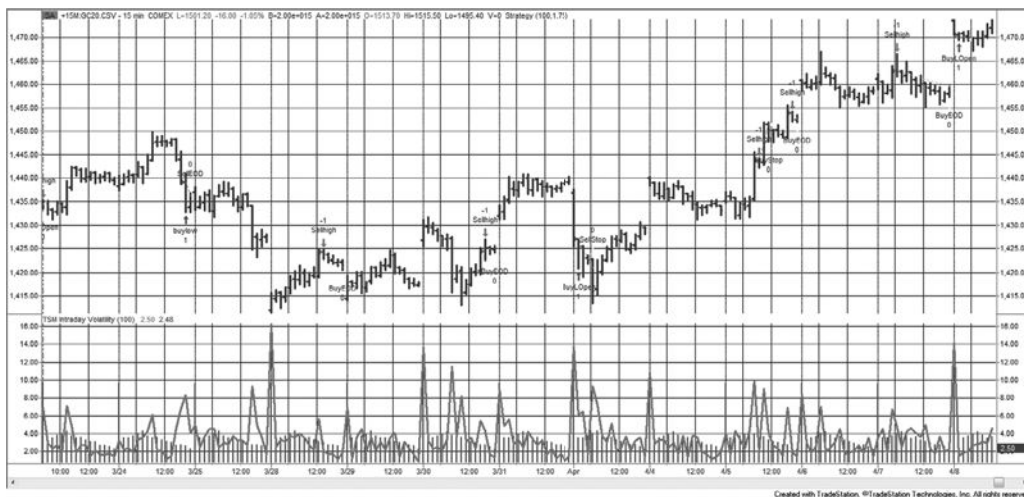
FIGURE 16.7 Intraday volatility for gold futures, April–May 2011, with trading signals.

3. If the bar is the open of the day, and today's true range is greater than the average of the previous bar (the last bar of the previous day) by a factor of 3, and

   a. The close is in the upper 20% of the range, then *sell short* on the open of the next bar

   b. The close is in the lower 20% of the range, then *buy* on the open of the next bar

   c. Save the high and low of the entry bar

4. If the bar is not the open of the day, then *buy* or *sell short* (3a or 3b) if the volatility of the current bar is 2 times the size of the average volatility of that bar and the close is in the top or bottom 20% of the bar range.

5. Do not place a stop for the next 2 bars (a delay to allow the volatility to continue), the place a *sell stop* below the low of the entry bar or following two bars, whichever is lower. For a short sale, place the *buy to cover stop* above the high of the three bars.

6. *Exit* on the close if not stopped out.

This is not a sophisticated strategy, but it does show profits for gold from 2000 using 20-minute bars, and for U.S. 30-year bonds using a volatility factor of 2.5. It has the basic ingredients of a reasonable strategy:

- The premise is that a larger-than-normal move will more often be caused by market noise and will tend to reverse.
- A stop loss in the event prices continue to move in the direction of the volatile bar (although delayed).
- An exit at the end of the day.

Some added rules might make this strategy more robust. For example, only those moves with higher volume (or lower volume) could be selected. Even more interesting,

we could adapt the method used by floor traders of watching the volume. While they actually listen to the level of screaming on the floor, the intraday volume can be a good substitute. Floor traders will buy on price moving higher as long as the floor noise was high or increasing. Once that noise started to fade, they exit the position, under the belief that whatever caused the excitement was over and price would pull back.

Using the volume chart in Figure 16.6, the strategy could exit when volume falls back to the average volume for that bar, or perhaps half the volume occurring on the high-volatility bar. That way the system is reading the market for its signals.

Adaptive Techniques

An *adaptive* technique is one that changes with market conditions. It could be as simple as using a percentage stop-loss; however, technicians associate the idea of adaptive with a broader scope—to handle more changing conditions without manual intervention. The main focus in this chapter will be on automatically changing the calculation period that determines the sensitivity and success of a trend or an indicator. Under some conditions, such as sideways markets, it may be best to use a slow trend or momentum indicator, and during explosive volatility those same methods may work better if they respond quickly. This shift from slow to fast (and back) due to changing market conditions is solved by using adaptive techniques.

ADAPTIVE TREND CALCULATIONS

We begin by looking at techniques that vary the length of the trend. These techniques are based on a common premise that it is better to use a longer-term trend, one that is slower to react to price change, when the market is either in a sideways pattern, exhibits low volatility, or, in the case of Kaufman, has high relative noise. This may seem a reasonable assumption, but identifying a sideways market is very difficult. A price may be unchanged from a week ago, yet during those 5 days it rose quickly for 3 days, then reversed sharply and, coincidently, was at the same level on the fifth day as it continued to plunge. In many ways these adaptive techniques improve the basic performance of standard trend calculations, but they cannot solve all of the problems.

Kaufman's Adaptive Moving Average

Kaufman's *Adaptive Moving Average* (KAMA),[1] which began taking form in 1972, is based on the concept that a noisy market requires a slower trend than one with less noise. *Noise*

---

[1]For a complete discussion of this method, see Perry J. Kaufman, *Smarter Trading* (New York: McGraw-Hill, 1995), 129–153.
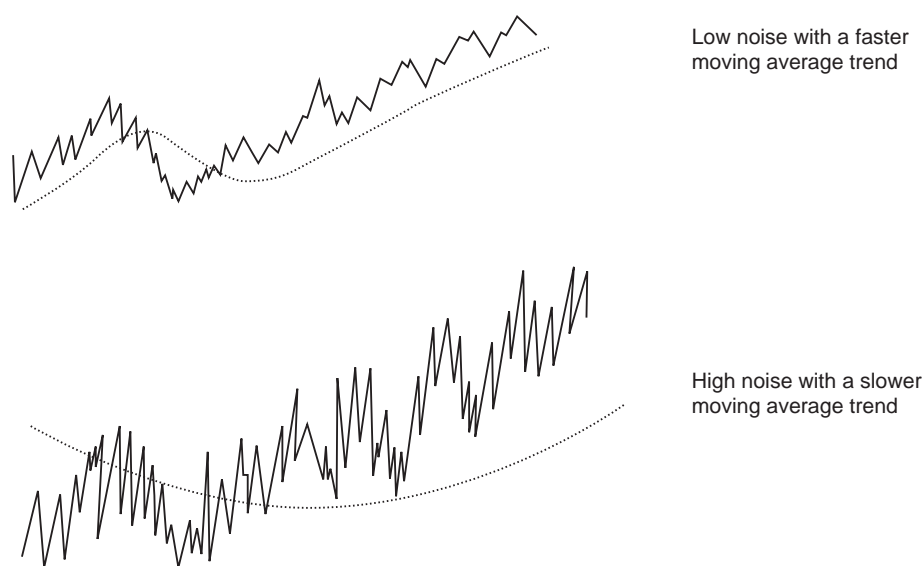
FIGURE 17.1 Similar price moves with low and high noise.

is the erratic price movement that can be seen on any chart as up-and-down price changes within a trend or during a sideways period. As we will show, noise is not just volatility. (For an explicit study of market noise, see Chapter 2.) The assumption in KAMA is that during a relatively noisy price move, the trendline must lag further behind to avoid being penetrated by the normal but erratic behavior of prices, which would cause an unwanted trend change. When prices move consistently in one direction with low noise, any trend speed may be used because there are no false changes of direction. Figure 17.1 shows a price move with low noise and the same move with higher noise. With low noise, the trendline can be positioned closer to the underlying price direction; with higher noise, the trend must lag farther behind in order to minimize penetration of the trendline.

KAMA is designed to use the fastest trend possible based on the smallest calculation period for the existing market conditions. It varies the speed of the trend by using an exponential smoothing formula (therefore, it is not a "moving average" as the name says), changing the smoothing constant each period. The use of a smoothing constant was selected because it allows for a full range of trends, represented as percentages, compared to the simple moving average, which uses integer values and is limited in its selection of the fast-end speeds (see the discussion of exponential smoothing in Chapter 7). KAMA is calculated as

$$KAMA_t = KAMA_{t-1} + sc_t \times (p_t - KAMA_{t-1})$$

where $KAMA_t$ = the new adaptive moving average value
         $KAMA_{t-1}$ = the previous adaptive moving average value
            p_t = the current closing price
          sc_t = the smoothing constant, calculated each period as follows:

$$sc_t = [ER_t \times (fastest - slowest) + slowest]^2$$

and

$$fastest = 2 \,/\, (fastest\ moving\ average\ period + 1)$$

$$slowest = 2 \,/\, (slowest\ moving\ average\ period + 1)$$

$$ER_t = \frac{\left| p_t - p_{t-n} \right|}{\displaystyle\sum_{i=t-n+1}^{t} \left| p_i - p_{i-1} \right|}$$

The *fastest* and *slowest* values represent the range of calculation periods over which KAMA can vary, each converted to its exponential smoothing constant equivalent. Nominally, these values are set to 2 and 30. By squaring the smoothing constant components, the slow end becomes 30×30, equal to an extremely unresponsive 900-period trend. The effect of this is to stop the movement of the trendline when the market becomes very noisy. On the other end, when the market goes up steadily, without a reversal for n days, the speed of the trend can reach a very fast 4-period equivalent.

The *efficiency ratio*, *ER*, is also now known as *fractal efficiency*. It has the value of 1 when prices move in the same direction for the full n periods, and a value of 0 when prices are unchanged over the n periods. When prices move in wide swings within the interval, the sum of the denominator becomes very large compared to the numerator, and *ER* approaches zero. Smaller values of *ER* result in a smaller smoothing constant and a slower trend. Both KAMA and the changing smoothing constant are shown in Figure 17.2. The 10-day KAMA is plotted in the upper part of the chart along with NASDAQ 100 prices during September 2003. The trendline moves steadily higher, turns down quickly, then begins up again. It appears to move in steps caused by the acceleration of the trendline, followed by its slowing down. The trendline never
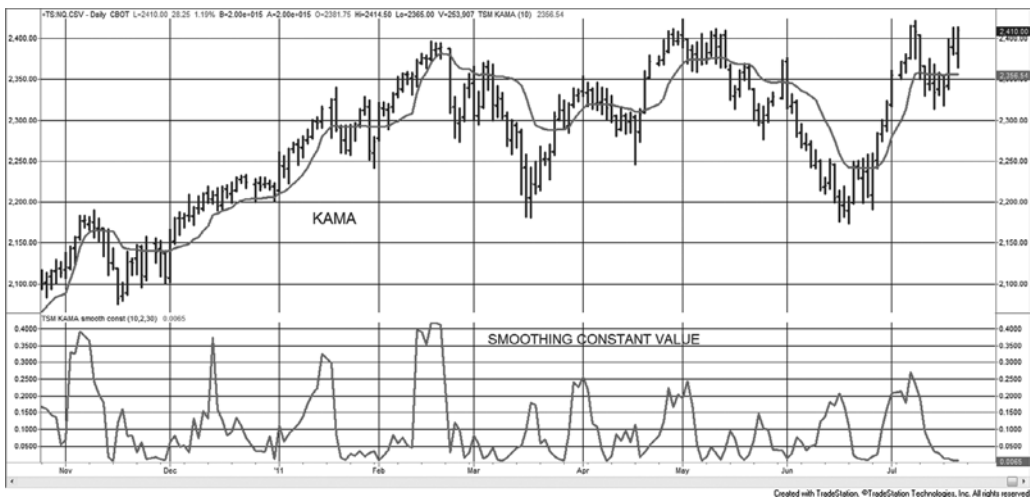
FIGURE 17.2 Kaufman's Adaptive Moving Average and the smoothing constant values on each day, *e*mini NASDAQ futures, November 2010 to July 2011.

retreats. Trading signals will be taken from the direction of the trendline, with some qualifications.

The lower part of Figure 17.2 shows the changing value of the smoothing constant, *sc*. In this example it reaches a high of just over 0.40, equivalent to a 5-day moving average, and a low near zero, which would cause the formula to assign the maximum 900-day calculation period. In the upper part of the chart, during January and May, the trendline appears to go sideways when the smoothing constant is very small.

If you choose to use the default values for KAMA, where *ER* is calculated over 10 periods, the smoothing constant ranges from 2^2 to 30^2, and C is the closing price; its calculation can be expressed in TradeStation code as

```
        KAMA = KAMA[1] + ((absvalue(C-C[10])/
  summation(absvalue(C-C[1]),10)* .6022) + .O645)*2*(C -KAMA[1])
```

where the bracketed numbers (e.g., [10]) represent *days back*.

Time Period for ER

In Kaufman's Adaptive Moving Average, the time period only affects the calculation of the efficiency ratio. The purpose is to keep that calculation period, n, as small as possible, so that the evaluation of only a few days will allow the speed of the trend to change from fast to slow and back again. Users may begin testing in the range of 8–10 days. It is not necessary to test calculation periods greater than the number of consecutive days that prices have moved in one direction. If prices rarely close more than 10 days in the same direction, then a 10-day *ER* is a good choice. Using more than 10 days will simply scale *ER* to a smaller number but not add any additional value to the calculation results.

Choice of Volatility

The use of the volatility function, *summation*(*absvalue*(*C-C*[1]),10) in TradeStation code adds all the individual close-to-close price changes rather than using the high-low range. The idea behind the method is to capture the efficiency of the price move. A very noisy move can be compared to a drunken sailor's walk, wandering from side to side (never backwards) yet still moving forward. Noise is not just volatility. If the Dow had a range of 100 points during the week, it matters whether it went up 20 points each day, or if it rose 100 points on Monday, dropped 100 on Tuesday, rose 100 on Wednesday, and so on. Clearly, going back and forth has more noise than a steady rise. Therefore, it is not as important that the price moved from a low point to a high point, but the pattern of how it was done.

Practical Matters

Occasionally there is a series of price moves that progresses in very small increments in one direction without a reversal, causing *ER* = 1. Although this is theoretically correct, it may prove best to include a filter for those cases where the average daily change for

some period (say 5 days) is less than a minimum threshold. Leaving the smoothing constant unchanged to avoid false signals may be the safest solution.

A second issue is caused by the use of exponential smoothing. While this technique is very convenient because changing the smoothing constant is all that is needed to change the speed of the trend, the negative feature is that the trendline always changes direction when it is penetrated by the price. This is counterproductive when trying to force the trendline to move sideways based on a very small smoothing constant. In Figure 17.2, there is a 2-week period at the far right where the KAMA trendline appears to go sideways while prices move above and below it. A very close look will show that the trendline actually changes direction at the same time prices penetrate the line, although these changes are very small. In the formula for KAMA, the expression $p_t - KAMA_{t-1}$ is where the new trend is positive when $p_t > KAMA_{t-1}$ and negative when $p_t < KAMA_{t-1}$.

To solve this, a very small filter is used to stop trend changes based on minor price penetrations. This filter can be a variable value, for example, 0.1 standard deviations of the trendline changes.

Trading the KAMA

The Adaptive Moving Average should be traded using the trendline to determine direction, buying when the trend turns up and selling when it turns down. Two other rules will be

- Keep the calculation period less than 14, 8 to 10 days is preferable. There are few occasions when prices move in the same direction 14 days in a row; therefore, any period longer than 14 will not change the values, only make them all smaller.
- Leave the upper bound of 30 days fixed. To make the trendline less sensitive, make the lower bound greater than 2.
- Use a small threshold to filter the changes in trend direction.

The Companion Website has one strategy, *TSM Adaptive Systems*, that allows testing of the KAMA as well as the Adaptive Correlation and VIDYA, which are discussed in the next sections. Additional indicators of the same names are available to plot these trendlines.

Fractal Adaptive Moving Average

John Ehlers has a number of variations on the Adaptive Moving Average, but one of the most interesting uses *fractals*.[2] Fractals are similar shapes (*self-similarity*) that can be pieced together to represent any contour or other figure. An important characteristic is that these shapes can be applied at any level of detail with similar results. The classic example is using a ruler to measure the shoreline of an island

---

[2]John F. Ehlers, "Fractal Adaptive Moving Average," *Technical Analysis of Stocks & Commodities* (October 2005).

(using a map). If we have a very large map and place a long ruler so that each end touches the coastline, we will get a smoothed value of the periphery of the island. The ruler will span some of the jagged parts of the coast. If we make the ruler smaller, the measurement of the coastline will increase because the ruler will include many of the irregularities skipped using a longer ruler. By following this logic, the coastline will be infinitely long if we use an infinitely small ruler. However, that is not a practical solution for trading.

Fractal dimension is a measure that describes the characteristic shape (or sparseness) of the figure (in our case a price series) at all levels of detail. The method of calculation is to cover the pattern with N number of the same patterns (*objects*) of various sizes, s. Then one solution might have 100 square boxes ($N = 100$) of 5 different sizes ($s = 5$).

According to Ehlers, the relationship of the number of objects in two sets of sizes is

$$\frac{N_2}{N_1} = \left(\frac{s_1}{s_2}\right)^D$$

$$D = \frac{\log\dfrac{N_2}{N_1}}{\log\dfrac{s_2}{s_1}}$$

As a simple example, we would like to find the fractal dimension of a line that is 10 feet long. We will place boxes of two sizes on the line, $s_1 = 1$ foot and $s_2 = 0.1$ foot. If we use 1-foot boxes, we can fit 10 boxes ($N_1 = 10$); if we use 0.1 foot boxes we can fit 100 boxes ($N_2 = 100$). Then the fractal dimension is

$$D = \frac{\log\left(\dfrac{100}{10}\right)}{\log\left(\dfrac{1}{0.1}\right)} = 1.0$$

Ehlers has used this to create an adaptive trend with features that include upper and lower bands, a trailing stop, and profit-taking. His original strategy has been modified and appears as the program *TSM AMA Fractal* (with a function and indicator) on the Companion Website.

Chande's Variable Index Dynamic Average (VIDYA)

Tushar Chande's *Variable Index Dynamic Average* (VIDYA)[3] first appeared in *Technical Analysis of Stocks & Commodities* in March 1992. Both VIDYA and Kaufman's KAMA use an exponential smoothing as a base for varying the speed of the trend each day. VIDYA starts with a fixed smoothing constant of 0.20 and varies that according to the relative volatility. Using two time periods, 9 and 30 days, we find the ratio of the

---

[3]Tushar S. Chande and Stanley Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994).

standard deviation of the closing prices. When the volatility is currently higher, the ratio is higher, resulting in an increase in the smoothing constant and subsequently a slower trend. Then higher volatility slows down the trend and lower volatility speeds it up.

$$VIDYA_t = k \times s \times C_t + (1 - k \times s) \times VIDYA_{t-1}$$

where s = a fixed smoothing constant, suggested as a 9-day equivalent
              (i.e., $s = 2 / (9 + 1) = .20$)
         C = the closing price
         t = the current period
     $t - 1$ = the previous period
         k = the relative volatility, calculated as

$$k = stdev(C,n) / stdev(C,m)$$

where $stdev$ = the standard deviation function
            n = the recent past n days and
            m = a longer historic interval in days, $m > n$

Chande uses a very practical relative volatility measure based on the standard deviation, or distribution of price moves over the past n days, compared to a longer historic pattern, taken over m days. When $k > 1$ the current volatility is greater than the historic volatility; when $k < 1$ the current volatility is lower. Typically, the standard deviation is taken using price changes, or returns, rather than actual prices. It is possible that the results might benefit from this change. VIDYA is included in the program *TSM Adaptive Systems* available on the Companion Website. An illustration and a comparison of its performance appear after the next section.

Correlation Coefficient, r^2

Another intuitive choice, suggested by Chande, uses the correlation coefficient, r^2. This value is the result of measuring the residuals from a linear regression calculation. When the regression fit is very strong, the value of r^2 is close to 1; when there is no apparent market direction, r^2 is near 0. This fits neatly into the pattern of smoothing constants for KAMA and VIDYA.

When programming this technique, the calculation of the correlation of the residuals is different from the standard correlation but can be done using the formulas found in Chapter 6. To avoid unnecessary complications, the following program calculates a similar technique (Method 1) by evaluating the correlation of the closing prices and a time series containing the sequential values 1, 2, 3, ... It then calculates the current value of the residual based on a linear regression of closing prices (Method 2) and finds the correlation coefficient of the series of current residuals. In the code for *TSM*

Adaptive Systems, Method 2 is plotted and the result of Method 1 appears in comment braces.

MESA Adaptive Moving Average

Ehlers uses the phase rate of change to create the *MESA Adaptive Moving Average* (MAMA), also fondly called the *Mother of Adaptive Moving Averages.*[4] The phase rate of change is 360° per cycle; therefore, if the price cycle is 36 bars in length, then the phase rate of change is 10° per bar. If there are only 10 bars in the cycle, then the phase rate of change is 36° per bar. The phase rate of change is the number of degrees in the cycle that is captured by the time it takes to produce a single bar. To find the phase rate of change, a solution similar to the Hilbert Transform is used (an explanation and example of the Hilbert Transform can be found in Chapter 11).

The calculation of MAMA uses the same basic exponential formula as KAMA; however, the value of the smoothing constant, *sc*, must remain between the *fast limit* of 0.5 and the *slow limit* of 0.05. The smoothing constant is

$$sc = fast\ limit\ /\ phase\ rate\ of\ change$$

where   if *sc* < 0 then *sc* = *fast limit*
          if *sc* > *slow limit* then *sc* = *slow limit*

Another indicator, the *Following Adaptive Moving Average* (FAMA), is created by applying the MAMA technique to the MAMA line itself, rather than to the price series, using a smoothing constant for FAMA that is half the value of the smoothing constant used for MAMA. The result is that FAMA is a smoother, less volatile version of MAMA. This serves as a *signal line* in the same way as the MACD or stochastic indicators. The crossover of MAMA and FAMA increases the reliability of the trend signal. A program to test MAMA is *TSM MAMA*, available on the Companion Website.

Comparison of Adaptive Methods

Figure 17.3 shows the two MESA adaptive trends (MAMA and FAMA are the dotted lines) along with KAMA, the adaptive correlation, and VIDYA, all applied to daily U.S. bonds futures from November 2010 through July 2011. During this period, there was one sustained downtrend and a similar uptrend. If we consider that a trading signal would have occurred on the cross of MAMA and FAMA, or when the other trendlines changed direction, then KAMA appears to be the slowest, followed by the correlation and MAMA methods, with VIDYA the most active.

The specific results for the four methods over 20 years of selected futures markets are shown in Table 17.1. Averages are along the bottom. It is interesting to note that

---

[4]John F. Ehlers, *Rocket Science for Traders* (New York: John Wiley & Sons, 2001).
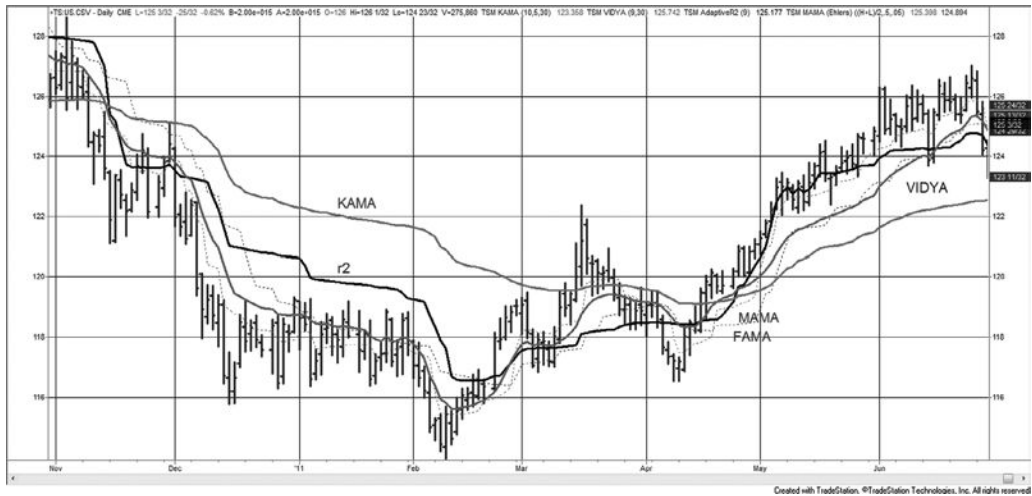
FIGURE 17.3 Four adaptive methods shown for 30-year bonds, November 2010 to July 2011. KAMA, Correlation (r^2), and VIDYA are the solid lines and MAMA and FAMA the dotted lines.

TABLE 17.1 Performance Comparison of KAMA, Correlation, VIDYA, and MAMA, 1990–July 2011 wherever possible.

	KAMA (per 10,5 to 30)			Correlation (per 20)			VIDYA (per 9,30)			MAMA (.05 to .50)		
	Profit Factor	Trades	Prof/ Ctr	Profit Factor	Trades	Prof/ Ctr	Profit Factor	Trades	Prof/ Ctr	Profit Factor	Trades	Prof/ Ctr
Eurodollars	2.27	143	308	2.00	230	182	1.64	376	97	2.35	173	286
U.S. 30-year bonds	1.31	175	189	1.16	215	150	1.14	455	87	1.18	206	170
EuroBund	1.83	145	415	1.59	179	301	1.33	386	120	1.28	188	172
emini S&P index	1.11	139	92	0.66	273	−272	0.70	548	−181	0.98	224	−12
NASDAQ index	1.97	94	798	1.06	174	57	1.01	359	9	1.53	140	455
Russell index	1.69	126	587	0.82	232	−180	0.89	418	−84	1.16	170	151
DAX	1.28	162	689	1.02	234	46	1.16	439	231	1.27	205	612
EURUSD FX	1.30	174	480	1.33	220	461	1.26	452	247	1.19	205	312
AUDUSD FX	1.14	188	110	0.90	239	−83	1.16	460	81	1.22	204	171
Crude oil	1.54	168	646	1.21	198	286	1.14	465	124	1.03	212	42
Heating oil	1.41	162	657	1.11	215	175	0.98	498	−18	1.14	220	212
Unleaded gas	1.41	193	565	1.13	221	213	1.10	471	116	1.27	217	429
Natural gas	2.09	158	1762	1.25	216	407	1.53	418	688	1.4	198	916
Gold	0.86	189	−118	0.92	203	−64	0.84	455	−90	1.16	222	112
Copper	1.26	173	306	1.49	178	566	1.12	455	85	1.61	220	547
Palladium	1.94	162	796	1.80	196	694	1.73	437	357	1.95	216	709
Average	**1.53**	**159**	**518**	**1.22**	**214**	**184**	**1.17**	**443**	**117**	**1.36**	**201**	**330**

each method is somewhat consistent in the number of trades, even though the markets are unrelated. For KAMA, 3-year bonds, EURUSD, and copper are all within 1 trade of each other. Success of the systems also seems to be inversely related to the average number of trades, as does the profits per contract, the order being KAMA, MAMA, r^2, and VIDYA.

One point of good news is that all methods generated profits on average over 20 years, and very few of the markets posted losses. That speaks well of this method, or trend-following in general. An average ratio of 1.53 for KAMA (even before costs), or 1.36 for MAMA, will translate into a very good portfolio return ratio.

It is difficult to compare the results of these methods with a standard moving average. Which calculation period would you use? Perhaps, because KAMA can vary from 25 to 900 (5 to 30 squared), a sample of all those periods would be the only way to measure results. When doing that, be sure that the distribution is progressive, such as a modified log or exponential selection, otherwise it will weight the sample toward the slow end.

ADAPTIVE VARIATIONS

The first part of this chapter dealt with adaptive trends using exponential smoothing as the means for changing the speed of the trend. The following sections show other approaches to self-adjusting trading techniques.

Making Momentum Calculations Adaptive

There are many momentum calculations, such as *RSI*, that vary from 0 to 1, 0 to 100, or from −1 to +1. In order to use any of these as a smoothing constant in an adaptive indicator, it is necessary to change the values to range between 0 and 1, where 0 is less trend and 1 is more trend, which is the normal interpretation for the RSI and most other indicators. If M represents the value of the indicator, then those values can be changed to fit the 0-to-1 pattern of the smoothing constant by using the transformations

Momentum values	Use the transformation
(1) M varies from 0 to 1	$sc = M$
(2) M varies from −1 to +1	$sc = \text{absvalue}(M)$
(3) M varies from 0 to 100	$sc = M/100$
(4) M varies from −100 to +100	$sc = \text{absvalue}(M/100)$

where sc = the smoothing constant to be used for the current period
M = the current momentum value

Transformation (2) recognizes that when the momentum is at +1 and −1 there is a stronger positive and negative trend. By taking the absolute values, the negatives
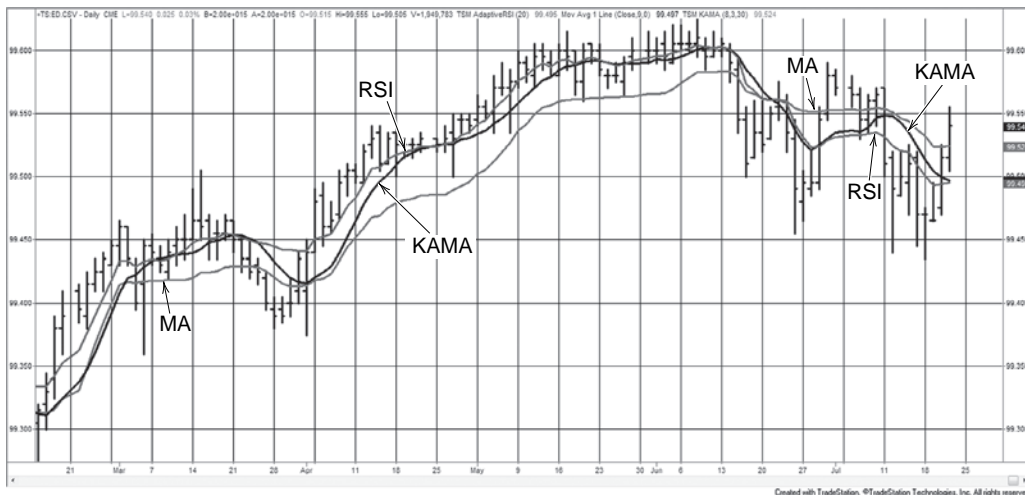
FIGURE 17.4 Comparison of the adaptive RSI, KAMA, and a 10-day moving average, applied to Eurodollar futures, February 2011–July 2011.

are reflected so that 0 to −1 take on the same pattern as the positive values from 0 to 1.

The RSI, changed to an adaptive indicator, is compared with the KAMA and a 10-day moving average, and applied to Eurodollar futures from February through July 2011, in Figure 17.4. The moving average tends to be the middle trendline. During the sideways periods, the adaptive RSI is more sensitive, the moving average next, and KAMA the least sensitive. This could be corrected in the adaptive RSI by squaring the smoothing constant. Then a value of 0.75 would become 0.56. Any lower value means more noise and a slower trend. A program to plot this is *TSM Adaptive RSI*, available on the Companion Website.

An adaptive stochastic can be created from the basic stochastic formula by changing the look-back number of days in a way similar to KAMA, based on market noise. In the following code, *period* is the 10 days over which the efficiency ratio is calculated, the *slowend* and *fastend* are the same as KAMA, and *per* gives a changing number of days over which to find the highest high and lowest low, the basis of the stochastic calculation. The raw stochastic (*K-fast*) is then smoothed over 3 days to get the *K-slow* version. This indicator can be found on the Companion Website as *TSM Adaptive Stochastic*.

```
efratio = effratiofreq(period)/100;
per = intportion(slowend - efratio*(slowend - fastend));
raw = (close - lowest(low,per))*100 / (highest(high,per) -
      lowest(low,per));
astoch = average(raw,3);
```

McGinley Dynamics

Among the work produced by John McGinley is a list of popular adaptive methods. His own *New McGinley Dynamics*,[5] is calculated as

$$MD_t = MD_{t-1} + (p_t - MD_{t-1}) / \left(k \times n \times \left(\frac{p_t}{MD_{t-1}} \right)^4 \right)$$

where MD_t = the current McGinley Dynamic
        MD_{t-1} = the previous McGinley Dynamic
        k = the constant 0.60, 60% of a selected moving average period n
        n = the moving average period
        p = the closing price

The Parabolic Time/Price System

The first well-known adaptive technique was the *Parabolic Time/Price System*,[6] which attempts to reduce the lag intrinsic to a trend system as the trend develops. To do this, Wilder increased the speed of the trend by shortening the number of days in the calculation whenever prices reached new profit levels. The philosophy of the Parabolic System is that time is an enemy. Once a position is entered, it must continue to be profitable or it will be liquidated.

The Parabolic Time/Price System is always in the market; whenever a position is closed out, it is also reversed. The point at which this occurs is called the *Stop and Reverse* (SAR). When plotted, the Stop and Reverse point seems similar to a trendline although it has a decreasing lag. This is accomplished by making the SAR point, the distance between the current price and the trendline, closer to the current price of each bar, as shown in Figure 17.5. During periods of short, consistent trending, the Parabolic SAR converges on the price trend, extracting excellent profits.

To calculate the SAR value, first assume a long or short position. If the market has recently moved lower and is now above the lows of that move, begin with a *long*. Call the lowest point of the previous trade the *SAR initial point* (SIP) because it will be the starting point for the SAR calculation (SAR_1 = SIP). Calculate each following SAR as

$$SAR_t = SAR_{t-1} + AF_t \times (High_t - SAR_{t-1})$$

This is an exponential smoothing formula using the high price for a long position and the low price when a short is held. For this method, the smoothing constant is called the *acceleration factor* (AF), and it is initially set to 0.02 at the beginning of each trade. After

---

[5] John McGinley writes *Technical Trends*, P.O. Box 792, Wilton, CT 06897, and has compiled a selection of adaptive techniques as well as developed the ones discussed here.

[6] J. Welles Wilder, Jr., *New Concepts in Technical Trading Systems* (Greensboro, NC: Trend Research, 1978).
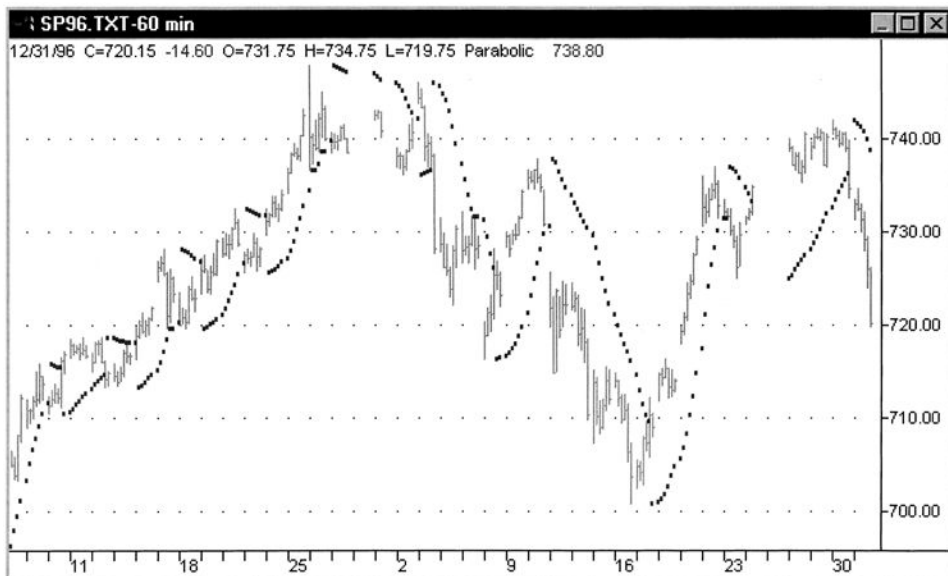
FIGURE 17.5   Parabolic Time/Price System applied to *e*mini S&P futures, February through July 2011.

a day in which a new extreme occurs (a new high when long, or a new low when short), the AF is increased by 0.02. In terms of moving average days, the AF begins at 99 days and increases speed to a maximum of a 9-day moving average, but not linearly. The acceleration factor, AF, cannot be increased above the value 0.20.

In the SAR calculation, the highest high of the current move is used when a long position is held, and lowest low when a short is held. This feature keeps the SAR at its highest possible level during an upwards move. An additional rule compensates for this strength and prevents premature reversal by not allowing the SAR to get any closer than the price range of the most recent 2 days:

> *If long, the SAR may never be greater than the low of today or the prior day. If it is greater than this low, set the SAR to that low value. A reversal will occur on a new intraday low which penetrates the SAR.*
>
> *For short positions, the SAR is initialized as the high of the recent move, AF = 0.02, and the daily calculation uses the lowest low of the current price move. To allow some price fluctuation, the SAR may never be below the high of today or the prior day.*

Wilder's Parabolic Time/Price System was unique in varying the smoothing constant to satisfy the notion that time worked against the trade. The idea of increasing the speed of the trend as the trade becomes more profitable is still a concept that has validity and has not yet been adopted in other adaptive methods. One weak point may be that AF always

begins at 0.02. Although this allows the trade latitude to develop without being stopped out, a market that is moving quickly through the buy or sell signal would be better traded with a faster trend. This method also requires fairly consistent price swings to produce profits.

A strong point is that the initial stop and reversal is the market extreme high or low rather than an arbitrary, computer-generated point. Also, the SAR does not get closer than the previous day's high or low, which prevents a reversal due to noise when the market is moving strongly in a profitable direction. For those readers interested in the further development of this method, the Parabolic Time/Price System is combined with Directional Movement (described in Chapter 23) to form the Directional Parabolic System discussed in Chapter 9.

The TradeStation function, available on the Companion Website as *TSM Parabolic*, allows the user to specify the starting value of the smoothing constant AF_{min}, the increment AF_{inc}, and the maximum value AF_{max}. This provides for more flexible testing. Optimization of these values has been reported to produce good results.

Alternate Calculation of Acceleration Factor

Volker Knapp suggests that the acceleration factor can be successfully replaced by a variation of Kaufman's *efficiency ratio (ER)*.[7] Using a 10-day ATR and a 10-day ER, the rules are:

1. Set the initial stop at 6 times the ATR from the entry price.
2. If *ER* < 0.30, then leave the stop unchanged.
3. If $0.30 < ER < 0.60$ then reduce the distance of the long stop by $0.1 \times ATR$ or the short stop by $0.5 \times ATR$.
4. If *ER* > 0.60, then reduce the distance of the long stop by $0.2 \times ATR$ or the short stop by $0.1 \times ATR$.
5. The trailing stop is only advanced when the new stop is higher for longs or lower for shorts.
6. The stop applies to trading on the next day.

Note that the rules are not symmetric for longs and shorts.

The Master Trading Formula

Another landmark method of creating a variable speed approach is Mart's *Master Trading Formula*.[8] It is also based on an exponential smoothing formula where the smoothing constant and band formed around the trendline are calculated daily, based on market volatility. Mart's general technique is to combine the *average true range* for the past

[7]Volker Knapp, "Parabolic Alternatives," *Active Trader Magazine* (September 2010).

[8]Donald Mart, *The Master Trading Formula* (Brightwaters, NY: Winsor Books, 1981).

15 days with the net change over the same time period. Each of these volatility elements is ranked from 1 to 21 by defining a maximum value for each one and dividing that maximum into 20 equal zones (the 21st being anything higher). The two rankings are then averaged to get a *correlated volatility factor* (CVF). Another ranking of 21 zones is created by dividing the highest CVF into 20 equal parts. The CVF then relates linearly to a range of smoothing constants from 0.084 to 0.330, representing the approximate exponential equivalents of the moving average days from 5 to 23. The trendline value is then created daily using the formula

$$Trend_t = Trend_{t-1} + CVF_t \times (HLC_t - Trend_{t-1})$$

where HLC_t = the average of today's high, low, and closing price

As with the other adaptive methods, the speed of the trend increases as volatility increases.

A band is also placed around the trendline, based on an inverse relationship to the correlated volatility factor. The band is widest when there is low volatility and the smoothing constant produces a slow trend; it is narrowest in fast, highly volatile markets. Rules for trading Mart's Master Formula follow the basic band trading rules. The system is always in the market, taking long positions when the upper band is broken and short positions when the lower band is crossed.

A number of observations can be made about the Master Trading Formula as compared to the previous adaptive methods. First is that the equivalent speed of the trend is the fastest, from 5 to 23 days compared to KAMA which ranges from 25 to 900 (5^2 to 30^2). The second is that the band gets narrower as volatility increases, while the trend slows, which is an attempt to enter faster on a penetration even while the trendline backs away. The combination would generate a signal sooner, which might be best for short-term, or intraday, trading. The last observation is that dividing the range into zones does not consider that there is less sensitivity in the calculation periods as the time increases. It might be best to have a non-linear distribution, where the zone increased as the trend slowed.

OTHER ADAPTIVE MOMENTUM CALCULATIONS

Varying the period of calculation does not have to be limited to a trend formula. The period for any technique can also be adaptive. The following techniques apply the same concept to traditional indicators that are used to measure overbought and oversold price movement.

Trend-Adjusted Oscillator

One problem that confronts many traders who use stochastics or other oscillators is that the oscillator values tend to cluster in the lower half of the chart when the trend is

down and in the upper half when the trend is up. For a countertrend trader, this causes many oversold conditions during a downtrend that can give false buy signals. A *Trend-Adjusted Oscillator*[9] (TAO) can correct this bias by modifying the value of the oscillator by the amount that the trend of the oscillator varies from its midpoint (50 for oscillators that range from 0 to 100, and 0 for those ranging from –100 to +100). Therefore, if a 10-day oscillator has been steadily declining and its 5-day moving average has a value of 40, then the current oscillator value is raised by 10% (the midpoint of 50 minus the oscillator moving average of 40). The calculations for this are:

$$n\text{-}Day\,oscillator\,(O_t) = \frac{\sum\limits_{i=t-n+1}^{t} \frac{close_i - low_i}{high_i - low_i}}{n}$$

$$n\text{-}Day\,moving\,average\,of\,oscillator\,(MAO_t) = \frac{\sum\limits_{i=t-n+1}^{t} O_i}{n}$$

$$Trend\text{-}adjusted\ oscillator\ (TAO_t) = Midpoint - (MAO_t - O_t)$$

Correcting the sustained overbought and oversold situations in an oscillator is an ambitious goal. The trend-adjusted oscillator succeeds in many ways but at the cost of losing some of the more extreme cases. In Figure 17.6, which applies the TAO to gold futures, the peaks and valleys are well-aligned through April 2011. When gold begins a more
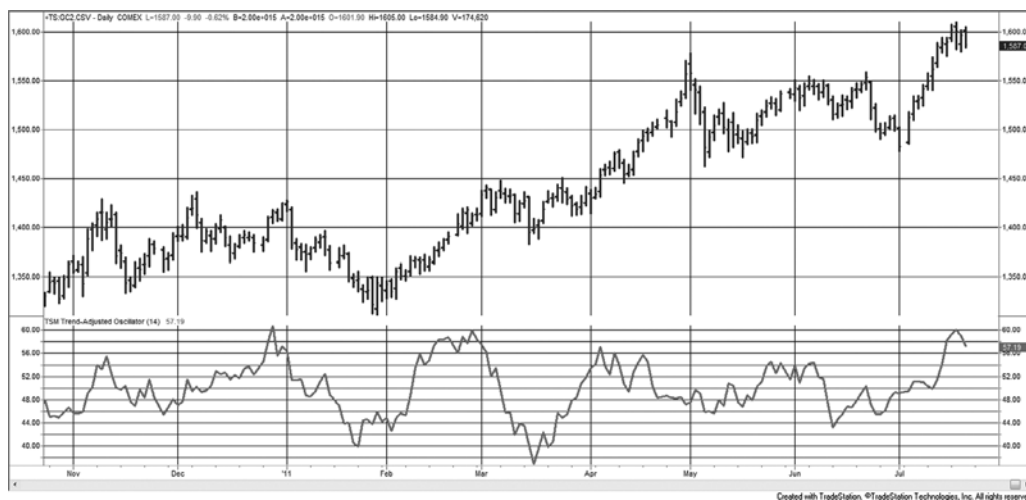
FIGURE 17.6   The Trend-Adjusted Oscillator applied to gold futures, November 2010 to July 2011.

---

[9]E. Marshall Wall, "Rolling with the Punches," *Futures* (July 1996).

sustained move up, the oscillator is gradually adjusted to compensate for the average value increasing toward the overbought level. This can be seen in the decreasing oscillator values in April while the prices continue to rise. At the peak of gold prices on May 1, the TAO is near 50, its neutral value, rather than at an overbought level. It takes until mid-July before the oscillator again adjusts to the price swings. While this fails to show the highest high, the normal stochastic would have been overbought from the beginning of April through May 1, also failing to isolate the high. This indicator is available on the Companion Website as *TSM Trend-Adjusted Oscillator*.

Dynamic Momentum Index

This approach to creating a dynamic, variable-length relative strength index, which changes the number of calculation days based on market volatility, can be used to change the period of any technique.[10] In the *Dynamic Momentum Index* (DMI, not to be confused with Wilder's *Directional Movement Indicator*), the length of the calculation period increases as the volatility declines. You may also select a pivotal period, around which the calculation period will vary; for the RSI this will be 14 days. When the volatility increases, the period will be less than 14 days, and when volatility declines, it will be greater. The calculations are

$$DMI_t = int\left(\frac{P}{V_t}\right)$$

$$V_t = \frac{\sigma_t(C,n)}{\left(\sum_{i=t-m+1}^{t}\sigma_t(C,n)\right)/m}$$

where V_t = volatility today, the standard deviation over n days divided by the average standard deviation over m days, where $m = 2 \times n$
     P = pivotal period, nominally 14
   $\sigma_t(c, n)$ = today's standard deviation, using the closing prices over n days
     int = integer function, returns only the integer portion
     n = days over which the numerator is calculated, nominally 5
     m = days over which the denominator is calculated, nominally 10

When the DMI is 14, the suggested pivotal value, the volatility is *normal.* That occurs when the current n-day standard deviation (the shorter) is the same as the m-day (the longer) average. When the current volatility increases, the ratio is greater than 1 and the value of the DMI becomes larger than 14.

A few alternatives in the DMI approach may improve results, and should be considered when generalizing this technique. First, the standard deviation should always be applied to the *changes* in price rather than the prices themselves. This removes the trend

---

[10]Tushar Chande and Stanley Kroll, *The New Market Technician* (New York: John Wiley & Sons, 1994).

component. Additionally, the volatility ratio may be more robust if the sum of the absolute value of the price changes is simply compared against the average of that sum, and that m is much larger than n rather than limited to a factor of 2,

$$Sum_t = \sum_{i=t-n+1}^{t} |C_i - C_{i-1}|$$

$$V_t = \frac{Sum_t}{\left(\sum_{i=t-m+1}^{i} Sum_i\right)/m}$$

This indicator is available on the Companion Website as *TSM Dynamic Momentum Indicator*.

ADAPTIVE INTRADAY BREAKOUT SYSTEM

An adaptive method developed for intraday stock trading, or stock index trading, has a number of unique features. The method is based on the idea that the expected value of the n-bar range is proportional to the square root of n, a concept derived from the classic price-volatility relationship. Taking a different approach, Meyers[11] first defines the high and low ranges as

$HR_t = H_t - L_{t-n}$ (today's high minus the low of n bars ago)

$LR_t = L_t - H_{t-n}$ (today's low minus the high of n bars ago)

Note that the high and low ranges *do not* use the highest high or lowest low of the past n bars, but the actual high and low of a single bar located n bars ago. Next, a range ratio is created from HR_t and LR_t by normalizing in two ways, (1) dividing by $ATR(n)$, the average true range over period n, and (2) dividing by a power function, n^a, to correct for the characteristic that the *range ratio* gets larger as n gets larger. These calculations allow the different lookback periods to be compared. The normalized ranges are then

$$NHR_t = HR_t / n^a / ATR(n)$$

$$NLR_t = LR_t / n^a / ATR(n)$$

Finally, the maximum range for *NHR* and *NLR* can be found over the most recent n periods,

$$HR_{max} = highest(NHR,n)$$

$$LR_{max} = highest(NLR,n)$$

---

[11] Dennis Meyers, "Range Roving," *Active Trader* (March 2003).

The final elements are *high* and *low threshold values* which will be used to generate trading signals.

<div align="center">

Buy when $HR_{max} > $ *threshold high* and $LR_{max} < $ *threshold low*

Sell when $LR_{max} > $ *threshold low* and $HR_{max} < $ *threshold high*

Close out all trades 5 minutes before the close of trading

</div>

Four key variables—n, the lookback period, a, the power function, and the two threshold values—are all found by optimization. It is suggested that the maximum lookback period n, should be no greater than 25. Tests performed on QQQ, the NASDAQ 100 ETF, yielded the values $n = 6$, $a = 0.75$, *threshold high* $= 0.50$, and *threshold low* $= 1.05$. If this method can be shown to be consistent over a reasonably long test period, producing about one trade per day, it is well worth pursuing.

AN ADAPTIVE PROCESS

An adaptive method can be a process as well as a formula. Rather than using an index or ratio to change the smoothing constant, which in turn alters the trend speed to be in tune with the current market, you can retest your system regularly using more recent data. The period being retested can always be a fixed number of days, or it can be selected visually, beginning when the market has changed its pattern, become more volatile, undergone a price shock, or moved to new highs or lows. The problem is having enough data to be satisfied that the results are dependable; the faster you react, the less data there is to make a decision. This approach, along with other testing methods, is covered in detail in the Chapter 21.

A Development Example

The size of price changes can also be used to vary the period of a moving average. When prices become more volatile, as measured by the standard deviation of the price changes, a shorter calculation period can be used to follow the market more closely. When there are smaller price changes, or stable volatility, the preference may be to slow the trend by increasing the calculation period.[12] The following steps can be used to build and personalize a variable-speed moving average with those characteristics:

1. Select the range over which the period of the moving average may fluctuate. For example, a medium-to-fast program might range from 5 to 30 days.

---

[12] George R. Arrington, "Building a Variable-Length Moving Average," *Technical Analysis of Stocks & Commodities* (June 1991).

2. Calculate the mean and standard deviation of price changes over a separate, fixed time period, recommended as the length of the longest trend period (in this case 30 days), but no less than 10 days so that there are enough days to have a reasonable sample. Using a smaller period reduces the response time needed to switch from one moving average speed to another, but causes the results to be more erratic.

3. Select standard deviation threshold levels, which separate price changes in a way that is seen to be relatively more active or less active. For example, we will define low volatility as falling inside the bounds created by the *mean* $\pm\, 0.25$ *standard deviations*. This will be a less active area where the longest calculation period would be used. Outside the bounds of the *mean* $\pm\, 1.75$ *standard deviations* would be a highly active market requiring a fast trend.

4. Establish the rate at which the moving average period will change as volatility (price change) increases from the inside to outside boundary. This could be the linear relationship

$$n_t = (30 - 5) \times \frac{\left(\dfrac{(p_t - p_{t-1})}{s} - 0.25 \right)}{(1.75 - 0.25)} + 5$$

$$\text{for } 1.75 > \frac{p_t - p_{t-1}}{s} > 0.25$$

where s is the standard deviation of the price changes (see step 2). When not between 1.75 and 0.25, n_t is the maximum and minimum values, 30 and 5, respectively. The general form is

today's moving average period = (max MA period – min MA period)
                    × (today's change / stdev of changes
                    – min stdev boundary) / (max stdev boundary
                    – min stdev boundary) + min MA period

CONSIDERING ADAPTIVE METHODS

At first glance, using an adaptive technique appears to be in conflict with the sound statistical approach that encourages the choice of a simple set of rules applied to a long period of test data. The classic result is a robust model, but one that might show considerable variation in performance over the test period. To stabilize returns in a market that has had enormous variations in volatility, we have moved away from using fixed values in our strategies, such as a \$500 stop-loss, or a 50-point breakout, substituting risk and entry criteria that are based on volatility. The simplest of these methods uses a percentage of price; the most complex can be very intricate functions of volatility and cycles.

When a variable feature is incorporated into the strategy, the values smoothly adjust to the market patterns. However, at each price or volatility level, in particular extreme high or low volatility, there are few occurrences; therefore, we are not as certain that the sliding scale works at all levels, or how it can work best at extremes. We can only judge the success of the technique by the improvement in the performance profile of the entire test period.

If we follow that logic further, we eventually come to the adaptive, or self-adjusting method, in which the most fundamental elements of a calculation can vary based on price level, volatility, or a broad choice of patterns. The methods in this chapter focused on trends and oscillators. The first varies the trend calculation period. The choice of a single trend does not serve us best for prolonged periods in the market. Slower trends are not reversed often, and give back much of their profits before ending; faster trends work only during periods when prices move quickly and have persistence. The obvious solution is that the calculation period should change, speeding up when there is a short-term, fast-moving price pattern. At other times, it should align itself with the long-term direction. This concept is perfectly reasonable, although formulating the rules to vary the period has not been perfected. The methods shown here may work better than a static period, but are far from the final solution.

Other adaptive techniques include adjusting the period of an oscillator, which is similar to the trend, and centering an oscillator to avoid prolonged periods where the value of the indicator loses significance by pushing against the upper or lower limits. By recording the high and low values of a past time interval, the oscillator values can be continually readjusted to provide a more useful tool. Analysts are encouraged to look further into these methods as an excellent way to improve system robustness.

Price Distribution Systems

Price movement is usually seen as a chart in which each new time period is a new bar or point recorded to the right of the previous prices—the traditional *time series*. There are many applications that need to look at prices differently. In options, the historic and implied volatilities are compared to decide the chances of prices remaining in a specific range for a specific amount of time. When a risk manager evaluates the volatility of performance, the same annualized standard deviation calculation is used. This was introduced in Chapter 2. The standard deviation gives the most basic and accepted measure of price distribution. From the standard deviation we can estimate the chances of a price remaining within a range over time. Because the standard deviation is the most commonly used measurement of price distribution, it is important to remember that a band formed by the average price change ±1 standard deviation contains 68% of the price movement (both up and down), ±2 standard deviations contain 95%, and ±3 standard deviations contain 99.7% of all price movement. Theoretically, all price movement should be contained within the range of ±3.5 standard deviations.

MEASURING DISTRIBUTION

The data used to determine the standard deviation is very important. Because we are seeking a statistical measure, it is most accurate when a large amount of data is applied. For example, the annualized volatility (using daily returns) of crude oil from 1983 through 2010 was 39%. But both historic and implied volatility are normally calculated and compared over 20-day periods. This allows you to see the changes in volatility. Figure 18.1 shows the cash prices of crude oil with the historic 20-day annualized volatility. At the end of 2008 volatility jumped to 140%. How is that possible? Volatility can spike for short periods because of extreme stress on the market causing very fast moves. In this case, after crude had risen from $30/bbl in 2003 to $150/bbl in 2008, it dropped back to $30/bbl in less than 6 months. While volatility could never be more than 100% of a downward price move, the standard deviation is reflecting upside volatility and does not know that
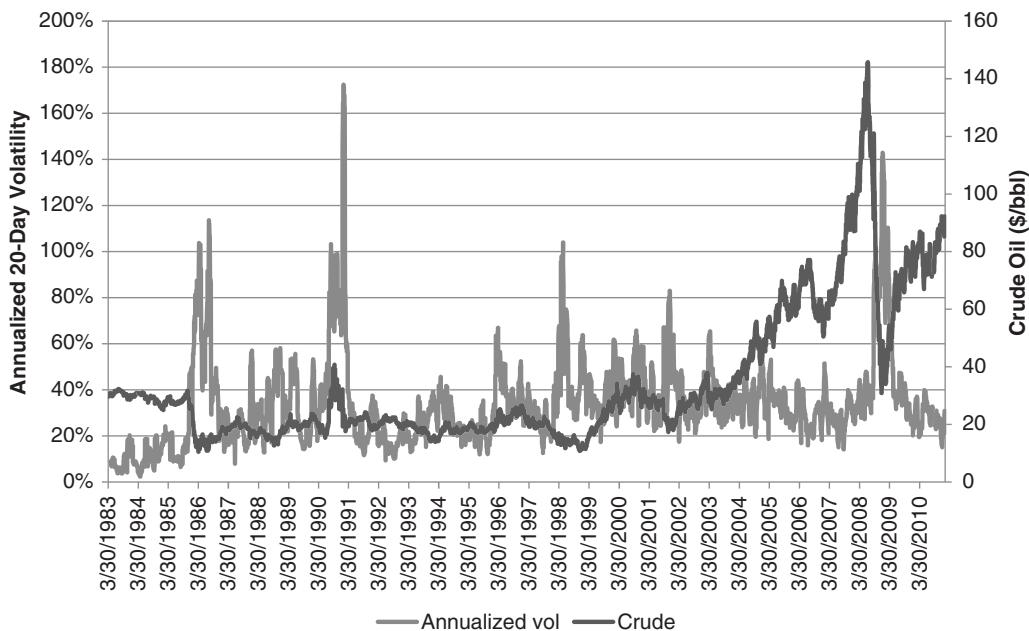
FIGURE 18.1 Crude oil prices and annualized volatility based on 20-day cash market returns, from April 1983 to February 2011.

prices cannot go below zero. Nevertheless, the standard deviation is the most useful tool that we have for measuring distribution.

The frequency distribution is another very practical approach to measuring price distributions. This was also covered in Chapter 2. It has the advantage of having a much clearer visual interpretation. By measuring a percentage from each end of the distribution, it adapts to skewness. While the standard deviation gives us what appears to be a highly mathematical probability, the large error factor that is caused by small amounts of data, as well as the inability to recognize asymmetry, reduces its usefulness. In the following sections, both techniques will be used.

Standard Deviation Bands

The use of *Bollinger bands*, discussed in Chapter 8, is a very popular application of basic price distribution. To create the standard Bollinger bands, calculate the standard deviation of prices (not returns) over a period of 20 days, and form bands of 2 standard deviations above and below the 20-day moving average (equivalent to the mean of the data). For those wanting bands that envelop more or fewer data points, the standard deviation factor can be larger or smaller. Signals produced with a larger band tend to be more reliable, much less frequent, and may have larger losses when prices turn the wrong way.

Once calculated, Bollinger bands can be displayed on any price chart, as seen in Figure 18.2. Because a 2-standard deviation band would normally enclose 95% of the

FIGURE 18.2 Comparison of 21-period and 65-period Bollinger bands, applied to the *e*mini S&P, July 2008 through August 2009.

data, a penetration of either the upper or lower band represents strength or weakness, or a less common outlier. Some traders will use a penetration to confirm a directional trade and others to set up a mean reversion trade.

Bollinger bands are most important for describing market volatility. A relatively narrow band translates into low volatility. By comparing a 21-period Bollinger band with a 65-period band, you can see the relative difference between shorter-term and longer-term market volatility in Figure 18.2. The thicker lines, based on the 65-period calculation, cross the short-term band at points that show relatively overbought and oversold situations, the decline in September 2008, and a smaller decline in February 2009. Because 21 and 65 days were used, we can think about this as a comparison between monthly and quarterly historic volatility, useful values for options traders.

Avoiding the Lag Problem

Applying any technique to a rolling time interval of the most recent n bars is a common method of keeping in tune with current market conditions. In the case of a simple moving average, we should be very familiar with the lag that is introduced when prices are trending; that is, when prices are moving steadily higher, the lag causes the trendline to fall behind.

There is a similar lag when using the most recent n bars to calculate a standard deviation, even when the data has been detrended. If we are measuring the volatility of the market, and prices rally quickly, the volatility rises. This will be seen as a larger value of 1 standard deviation measured over a fixed number of days or bars. As volatility increases, the bands increase in width slowly. If we are using a 20-day

calculation, then today's increased volatility will only be ¹⁄₂₀ of the total. That will make the bands increase in width in an orderly manner. At the two major price turns in Figure 18.2, in October 2008 and March 2009, the band width increases as the price move continues in the new direction. In Figure 18.3, showing Intel, we can see a single Bollinger band continuing to increase after the bottom in October 2002 and after the top in December 2002.

The same lag occurs when volatility declines. The difference is that the bands remain wide and narrow slowly. This can be a problem if you want to measure another increase in volatility following shortly after the first. The bands are slow to correct. If the price move was upwards, as seen in the center of the chart, then the bulging bands that follow the peak make it more difficult for a short-term price drop to penetrate the lower band.

If the purpose is to identify unusual volatility, then it is better to compare the current volatility with a longer-term measurement, a lagged measurement, or both. The longer-term period gives a better measure of normal volatility and makes any new volatile period a small part of the total, so that it has less impact. The intention of lagging the volatility measurement is to avoid using the current price data as part of the calculation. Then if you estimate that a volatile price period lasts for one month, use a 6-month volatility ending 1 month ago, or a 12-month volatility ending 3 months ago. Then two volatile periods, one after the other, will both be recognized.
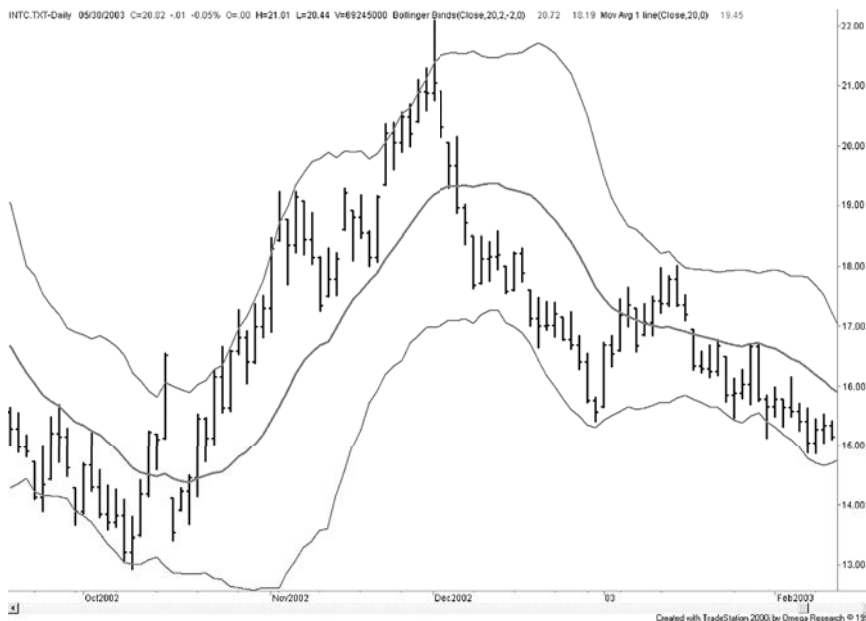
FIGURE 18.3   Standard deviation bands, measured from a trendline, show significant lag after a volatile price move.

USE OF PRICE DISTRIBUTIONS AND PATTERNS TO ANTICIPATE MOVES

Prices often form patterns that can be evaluated using probability methods, or simply viewed in much the same way as a frequency distribution, or histogram. Because many concepts are sound, but the statistical analysis is often difficult due to limited amounts of data or changing conditions, analysts have taken a much more empirical approach. The following section looks at some innovative ways to look at price distributions and how they are interpreted into trading opportunities.

Analysis of Zones

Rather than using standard deviations to identify the chance of a price move above or below yesterday's closing level, Bruce Gould observed that historic prices could be divided into five zones, each 20% of the price range over the previous 3 years. Using this long-term approach, it is easy to see that selling in Zone 1 (the lowest price levels) would have less opportunity for profit than selling in Zone 2, just above it. Similarly, buying in Zone 5, the highest band, would carry both the greatest risk and the least opportunity of profit. This is likely to remain the case unless prices move to much higher levels and all zones need readjusting.

J. T. Jackson[1] has used this concept to define 5 short-term zones, based only on yesterday's prices, which can be associated with the strength or weakness of today's move, and traded as an intraday strategy. These daily zones, which are popular with floor traders, are calculated as:

Calculation	Zone	If Price Is Above, Then
High2 = Average + High1 − Low1	6	Strong up
High1 = 2*Average + Low	5	Moderate up
Average = (High + Low + Close) / 3	4	Mildly up
Low1 = 2*Average − High	3	Mildly down
Low2 = Average + Low1 − High1	2	Moderately down
	1	Strongly down

where the High, Low, and Close are yesterday's prices. Note that there are 5 calculations but 6 zones needed to separate them. A test of how the S&P 500 falls into these relative rankings gives:

Zone	6	5	4	3	2	1
S&P Frequency	20%	44%	83%	79%	42%	20%

---

[1] J. T. Jackson, *Detecting High-Profit Day Trades in the Futures Markets* (Brightwaters, NY: Windsor Books, 1994).

which shows the slightly upward bias expected of the overall equities market. These distributions may vary depending on the interval used in their calculation and whether there is a dominant trend during that period. A longer interval that includes bull, bear, and sideways markets would be safest; otherwise, there is the chance that the calculations will create a bull market profile, which would not serve well during a bear market reaction. Although you can avoid trend bias by using longer intervals for the calculations, the zones tend to get very large.

The strategies for trading price zones focus on short-term trends and holding periods. For example, you can sell short when prices move into Zone 4 (mildly up) with a stop in Zone 5. If you consider Zones 3, 4, and 5 as containing mostly market noise, then selling at the top of Zone 4 and closing out that trade at about the average, or buying near the bottom of Zone 3 and closing out at about the average, could capture the majority of price moves that have no direction.

A program for testing this method is *TSM Zones*, available on the Companion Website. It requires both 10-minute, intraday data, and daily data.

Non-Random Patterns

In evaluating the zone approach, the markets that offer the greatest potential for this strategy are those that show an abnormal distribution of prices within the six zones. For example, if the six zones were all equal in size, and the frequency of prices declined by one-half as they moved from the center to the extremes, there would be a perfectly random distribution and no profit potential:

Zone of Equal Size	6	5	4	3	2	1
Normal Distribution	7%	14%	28%	28%	14%	7%

If the distribution is normal, but Zones 3 and 4 are much wider than the outer zones, then the risk of selling at the top of Zone 4 with a stop at the top of Zone 5 and a profit target at the average of Zones 3 and 4 would result in an equal number of profits and losses, but the profits would be larger. If the zones are of equal size, the opportunities come when the distribution is clustered in the center:

Zone of Equal Size	6	5	4	3	2	1
Normal Distribution	2%	8%	40%	40%	8%	2%

In this case, selling at the top of Zone 4 would result in many more profits than losses, although profits and losses would be of equal size and there would be fewer opportunities.

Applying a Moving Average Distribution

To extend the time frame includes a trend bias within the zone values, the same zones may be created by using the moving average values applied separately to the close, high, and low. For example, if three separate 21-day moving averages are calculated on the high, low, and closing prices, the moving average values can be substituted for yesterday's high, low, and close to give new zones. In using this approach, a strong upward trend

would cause all the zones to lag below today's prices, making current values strongly overbought. Zones created from the history of these averages will reflect the relative overbought and oversold price levels within trends. In a strong upward market, prices can remain in Zone 6 while they are steadily moving higher, in much the same way that an oscillator can remain over 80% for a sustained period.

Using moving average values will change the way in which you trade these zones. For example, you might want to enter long positions only in Zone 5 and look to exit in Zone 6. You might find that, if prices end the day in Zone 6 (very strong), they are likely to open the next day in Zone 4 (slightly higher).

Zones for Forecasting Range and Controlling Risk

Statistics have proved that, barring a superior forecasting method, the best estimate for tomorrow's price is today's price. That is, under most conditions, we cannot predict with any certainty that prices will go up or down tomorrow; therefore, the best estimate is to say that prices will be unchanged. However, if a trend system, such as a moving average, has been profitable, then its forecast for tomorrow is better than today's price. Market volatility, based on price changes, can be used with a directional forecast of tomorrow's price to create a set of zones used to control risk or project the probable trading range.[2]

First calculate the 10-day moving average of the absolute value of the daily price changes,

$$A_t = \frac{\sum_{i=t-n+1}^{t} |C_i - C_{i-1}|}{n}$$

where $n = 10$

Taken as positive numbers, zones are created that center around the current price and expand according to the average price change (volatility) using the following calculations:

$$H2 = C_{t-1} + 2A_t$$

$$H1 = C_{t-1} + A_t$$

$$L1 = C_{t-1} - A_t$$

$$L2 = C_{t-1} - 2A_t$$

Five zones are then created by the areas above *H2* and below *L2*, and the three ranges between *H2*, *H1*, *L1*, and *L2*, all of which change in proportion to the *n*-day volatility. These volatility levels, or zones, represent a very similar scenario to channel

---

[2] Based on Tushar Chande and Stanley Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994), 172.
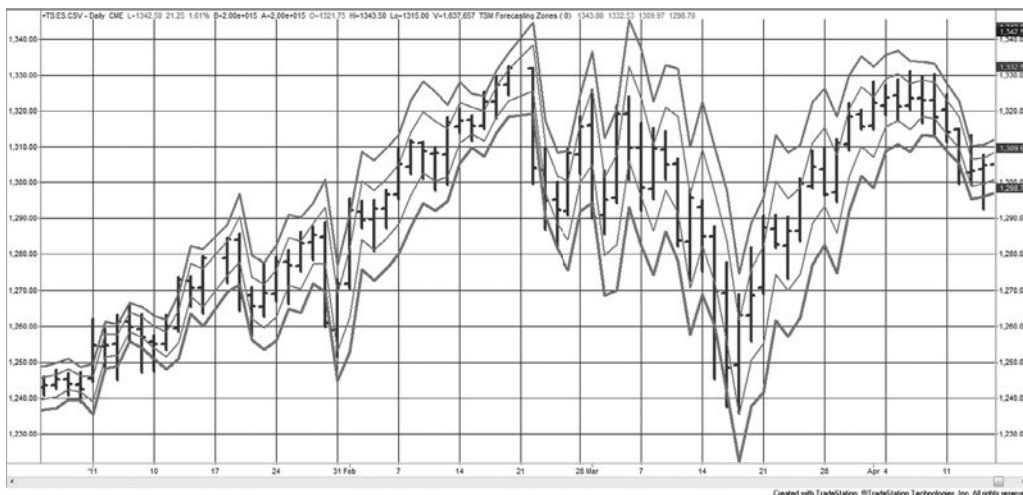
FIGURE 18.4 Forecasted trading zones using the Chande and Kroll method applied to the *e*mini S&P January–April 2011.

breakouts. The market often trades in a range defined by a normal or average level of volatility. When a new piece of information affects the price, it jumps to a new level, then trades with similar volatility (slightly higher at first) at the new level. Most often, the first breakout of an existing trading range puts prices in a zone just above the old range, making the pattern appear to be divided into equal zones. This same philosophy is the reason that standard profit targets for a price breakout are equal to the previous trading range. Figure 18.4 shows the forecasted zones applied to the *e*mini S&P, daily data, from January through April 2011. This was produced by the program *TSM Forecasted Zones*, as well as the spreadsheet of the same name, available on the Companion Website. Readers can find additional trading range projections in Chapter 15.

Zones Based on Daily Price Levels

Most traders watch the previous day's high, low, and close as key levels; therefore, it makes sense to define zones based on those values.[3] If we define six zones as

Zone 5:	Above today's high + 10-day average true range
Zone 4:	Above today's high and below zone 5
Zone 3:	Above (high + low)/2 and below zone 4
Zone 2:	Above the low and below zone 3
Zone 1:	Above the low − 10-day average true range and below zone 2
Zone 0:	Below Zone 1

_____

[3]Adapted from Massimiliano Scorpio, "Targeting Your Pattern," *Technical Analysis of Stocks & Commodities* (July 2005).

then we can look for two patterns:

Pattern 1:   Tomorrow's opening zone compared to tomorrow's closing zone
Pattern 2:   Today's closing zone compared to tomorrow's opening zone

Two different sessions for S&P futures from 2000 through 2011 are compared to show why 24-hour data (the electronic session) gives a different picture with fewer gaps. It is not necessary to use the day session as it exists. You can create your own "day session" using intraday data by specifying the start and end times. Table 18.1 shows the results

TABLE 18.1   Average Results of Patterns Using Tomorrow's Open and Tomorrow's Close

Open to Close		*emini* (Globex) S&P				Day Session S&P			
Entry Zone	Exit Zone	Count	% Count	Pattern Count	% of Pattern	Count	%Count	Pattern Count	% of Pattern
0	1	19	0.65		28.4	33	1.12		31.1
0	2	22	0.75		32.8	28	0.95		26.4
0	3	20	0.68		29.9	30	1.02		28.3
0	4	6	0.20		9.0	15	0.51		14.2
0	5	0	0.00	67	0.0	0	0.00	106	0.0
1	1	13	0.44		43.3	194	6.61		52.3
1	2	12	0.41		40.0	98	3.34		26.4
1	3	2	0.07		6.7	55	1.87		14.8
1	4	3	0.10		10.0	24	0.82		6.5
1	5	0	0.00	30	0.0	0	0.00	371	0.0
2	1	388	13.22		30.7	274	9.34		29.9
2	2	412	14.03		32.6	289	9.85		31.5
2	3	287	9.78		22.7	233	7.94		25.4
2	4	168	5.72		13.3	117	3.99		12.8
2	5	7	0.24	1262	0.6	4	0.14	917	0.4
3	1	215	7.32		13.9	127	4.33		12.3
3	2	246	8.38		15.9	195	6.64		18.9
3	3	498	16.96		32.2	339	11.55		32.8
3	4	552	18.80		35.7	353	12.03		34.2
3	5	37	1.26	1548	2.4	18	0.61	1032	1.7
4	1	3	0.10		11.1	27	0.92		5.4
4	2	1	0.03		3.7	48	1.64		9.5
4	3	9	0.31		33.3	100	3.41		19.9
4	4	9	0.31		33.3	277	9.44		55.1
4	5	5	0.17	27	18.5	51	1.74	503	10.1
5	1	0	0.00		0.0	0	0.00		0.0
5	2	1	0.03		50.0	1	0.03		16.7
5	3	0	0.00		0.0	0	0.00		0.0
5	4	0	0.00		0.0	3	0.10		50.0
5	5	1	0.03	2	50.0	2	0.07	6	33.3

of all start and end combinations of the open-close pattern on the same day. For the electronic market, most starting values fall within the high-low range of the previous day (Entry Zones 2 and 3), given that there is only a small time gap between the previous day's close and the next open. For the day session, which trades from 9:30 A.M. to 4:15 P.M., New York, there are many more starting values above the previous high (Zones 0 and 1) and below the previous lows (Zones 4 and 5), but a clear bias to the upside.

Table 18.2 is an overnight trade, entering on the close of the same day and exiting on the next open. There are many more cases were prices closed in the highest zone and

TABLE 18.2 Comparison of Patterns for the emini S&P and the Day Session S&P

Entry Zone	Exit Zone	emini (Globex) S&P				Day Session S&P			
		Count	%Count	Pattern Count	% of Pattern	Count	%Count	Pattern Count	% of Pattern
0	1	3	0.1022		0.0	25	0.8518		0.0
0	2	63	2.1458		0.0	69	2.3509		0.0
0	3	1	0.0341		0.0	6	0.2044		0.0
0	4	0	0		0.0	0	0		0.0
0	5	0	0	67	0.0	0	0	100	0.0
1	1	19	0.6471		3.1	138	4.7019		0.0
1	2	555	18.9033		89.7	314	10.6985		0.0
1	3	43	1.4646		6.9	138	4.7019		0.0
1	4	2	0.0681		0.3	34	1.1584		0.0
1	5	0	0	619	0.0	0	0	624	0.0
2	1	6	0.2044		0.9	125	4.2589		19.9
2	2	346	11.7847		51.6	197	6.7121		31.3
2	3	315	10.7289		46.9	197	6.7121		31.3
2	4	4	0.1362		0.6	108	3.6797		17.2
2	5	0	0	671	0.0	2	0.0681	629	0.3
3	1	5	0.1703		0.6	99	3.3731		13.6
3	2	298	10.1499		37.4	197	6.7121		27.1
3	3	482	16.4169		60.6	260	8.8586		35.8
3	4	10	0.3406		1.3	169	5.7581		23.3
3	5	1	0.0341	796	0.1	1	0.0341	726	0.1
4	1	0	0		0.0	42	1.431		0.0
4	2	45	1.5327		5.8	164	5.5877		0.0
4	3	724	24.6594		92.7	441	15.0256		0.0
4	4	12	0.4087		1.5	196	6.678		0.0
4	5	0	0	781	0.0	2	0.0681	845	0.0
5	1	0	0		0.0	0	0		0.0
5	2	0	0		0.0	0	0		0.0
5	3	0	0		0.0	0	0		0.0
5	4	0	0		0.0	0	0		0.0
5	5	1	0.0341	1	100.0	1	0.0341	1	100.0

those in the zones between the previous high and low are more evenly distributed with regard to exits. The program to produce this study is *TSM Zone Trades*, available on the Companion Website.

DISTRIBUTION OF PRICES

In the search to understand how prices move and what to expect, an analysis of price distributions can explain whether the market is trending, moving sideways, or simply unstable. Some of these patterns are clear, and others are not; in addition, the combination of patterns within patterns can become complex. The goal is to find a statistical edge. The following is intended as a basic approach to interpreting price distributions, although each group of markets has special characteristics. Before engineering a systematic approach to trading, it is best to understand how prices are expected to distribute. This approach avoids surprises and the risks that goes with them.

Long-Term Price Distributions

A quick observation of price data for most commodities, such as soybeans or gold, shows that we should expect a skewed long-term price distribution, with prices clustering at lower levels and a long tail representing higher prices. But even this basic fundamental rule has gotten more complex in the past 10 to 15 years. Gold has gone much higher and shows no signs of falling back to previous levels. Even wheat has moved out of its old range. Both of those markets, and commodities in general, are priced on the world market. Each country views those costs in terms of its own currency. And while gold is near $1600/oz at this time, the U.S. dollar had declined to 1.45 against the euro, so in adjusted terms, gold would be at $1100/oz. Then, does wheat still conform to the old distribution? Yes, but only if you remove the changing value of the U.S. dollar. Then a buyer in Europe or Asia will see the distribution of wheat prices as normal, while a U.S. buyer will see the price rising. To visualize this, Figure 18.5 shows the cash price of wheat (converted to an index), the same cash price deflated by the change in the U.S. dollar, and deflated separately by the change in the PPI.

The chart shows that the price of wheat is now 170% of the price at the end of 1985. Then U.S. buyers appear to be paying much more than in the past. If we remove the change in the PPI and the drop in value of the U.S. dollar, we see that real prices are actually less than they were in 1985 (below 90). Then the production of wheat has kept up with demand, but the U.S. dollar and inflation have pushed prices steadily higher. If we view the distribution based on actual cash prices (Figure 18.6a), the frequency distribution shows a longer tail on the left as prices move higher and never return. By removing the change in the U.S. dollar, the frequency distribution (Figure 18.6b) changes back to the classic form, with the long tail to the right.

When working with long-term distributions, it may be necessary to reduce the data by factors that have materially affected prices over the years. There is usually a simple way to decompose prices so that they show their true patterns.
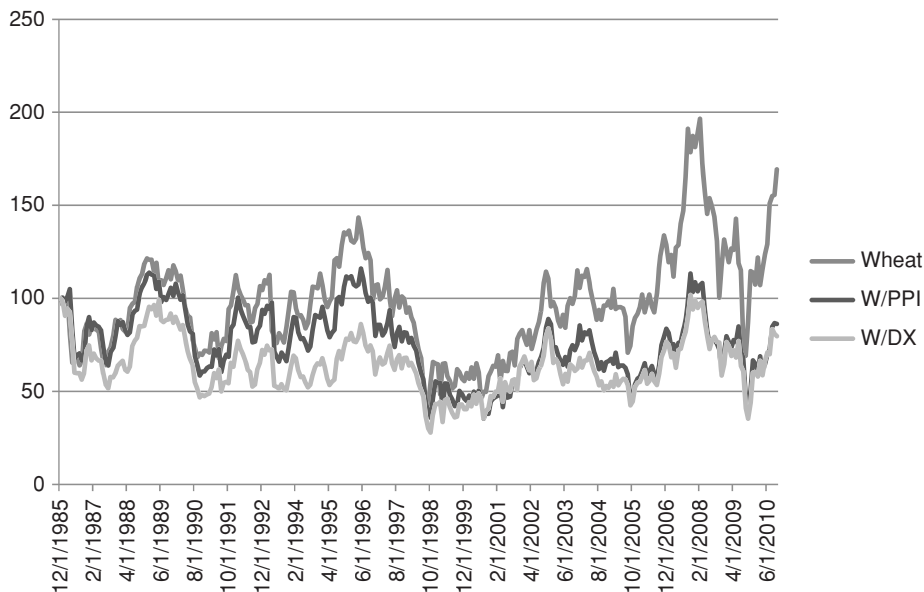
FIGURE 18.5   Wheat cash prices compared to the wheat price deflated by the dollar index and the PPI, 1985–2010.

The skewness in distributions means that the standard deviation is not a good measure. For any significant positive or negative skew (the tail to the right or left), the standard deviation is distorted by the distribution tail. If the tail is to the right, and the mean near the center, then 3 standard deviations to the left is likely to be off the chart, which makes it impossible to assign a probability to negative returns (or values that are less than the average).
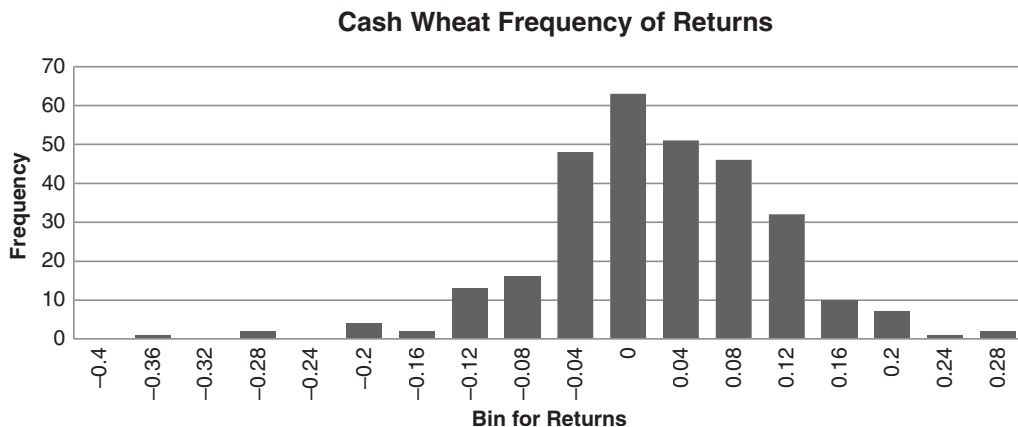
FIGURE 18.6a   Cash wheat frequency distribution of returns shows the tail to the left.

Wheat x DX Frequency
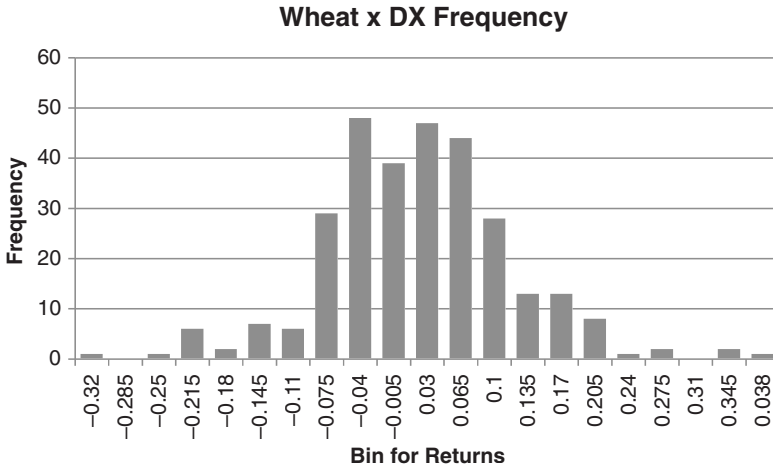
FIGURE 18.6b   By removing the changing value of the U.S. dollar using the dollar index, the skewness changes.

A large positive skew is particularly prevalent in the returns of a trend-following system. Because a trend system captures the fat tail, the very large, unexpected profits, and moving averages cut losses short, the returns are very skewed to the right. For example, if we use an 80-day moving average applied to U.S. 30-year bonds for 20 years, we get the distribution shown in Figure 18.7. The mean return is $232 and 1 standard deviation is ±$2621. Then 3 standard deviations are $8097 and –$7632. Those values fall in the middle of the right tail and well off the chart to the left.
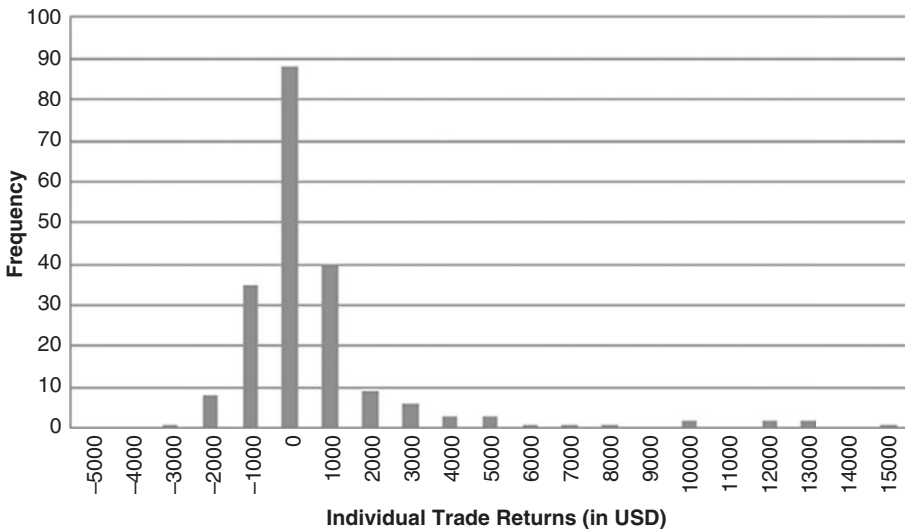
FIGURE 18.7   Frequency distribution of individual trade returns, U.S. 30-year bonds, 1990 through July 2011. Results are skewed to the right.

If you want to know the highest and lowest 10% for a skewed distribution, the most practical solution is simply to sort the numbers from high to low (the returns or P/L), then count from each end the equivalent of 10% of the data. For the PLs in Figure 18.7, 10% from the bottom and top would give –$1625 and +$2062, both quite far from the ends of the distribution.

Equilibrium

Most markets have periods of volatility that move sharply higher and lower. At some point, that volatility drops, and prices move within a narrow range. We might consider it a point of equilibrium. Low volatility often means that prices have reached a level that is generally acceptable to both buyers and sellers. For commodity markets, the equilibrium level tends to be at the lower end of the price range (adjusted for inflation or the U.S. dollar), but for currencies the equilibrium level is whatever is satisfactory to the governments. Any move away from equilibrium, either higher or lower, usually increases volatility.

Adjusting for Inflation

When trading, it is normally better to inflate the older prices so that the current price remains the same. In that way, any trading decisions will show the correct prices rather than needing to adjust them by the amount added or removed due to inflation or the changing U.S. dollar. Table 18.3 is an example of inflating monthly wheat cash prices by the dollar index (DX). The fourth column shows the monthly returns of DX, and the rightmost column is the inflated wheat price. The last row of the inflated price (October 2010) will always be the same price as the current wheat price.

Starting with the last price, 713, reverse the effect of the changing DX using the prior monthly value. Because DX dropped 1.98% in October, the September price is higher by 1.98%, using the formula E8=E9*(1 – D9), where column E is the inflated price and row 9 is October. Note that this is the opposite of the typical compounding rate-of-return formula. Of course, changes would be much smaller if daily, rather than monthly, data was used.

TABLE 18.3 Inflating Historic Wheat Prices to Remove Changes in the U.S. Dollar

Month	Wheat	DX	DX Returns	Wheat Inflated by DX
3/31/2010	383	82.77	0.0073	751.08
4/30/2010	416	83.20	0.0053	755.04
5/28/2010	452	87.93	0.0569	800.56
6/30/2010	477	87.18	–0.0086	793.74
7/30/2010	594	82.50	–0.0537	753.30
8/31/2010	618	83.95	0.0176	766.78
9/30/2010	622	79.37	–0.0545	727.15
10/29/2010	713	77.80	–0.0198	713.00
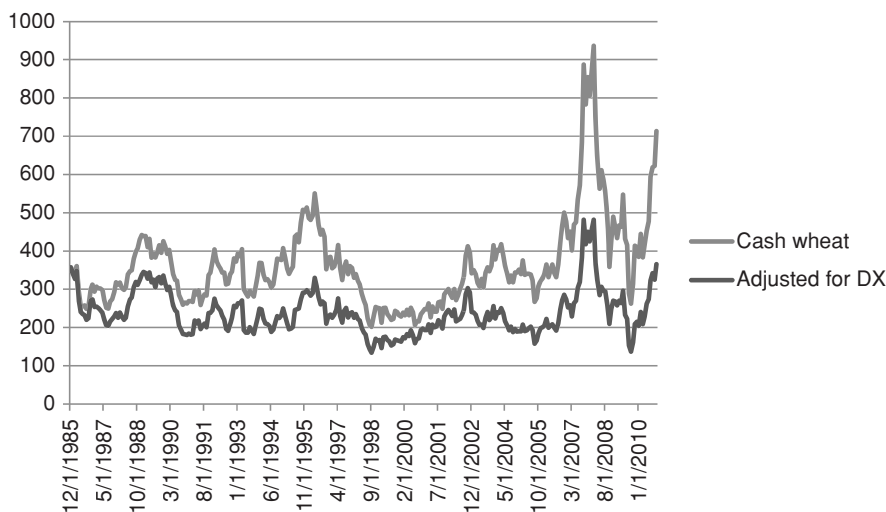
FIGURE 18.8   Wheat prices with historic data inflated by the value of the dollar index. Most recent prices are the same.

Another simple way of adjusting prices is to index the DX series so that it starts at 100, then multiply the wheat cash price by the indexed DX value. That will deflate the wheat prices so we can see how they changed over time. Figure 18.8 shows the original cash series and the deflated series going back to the end of December 1985. Prices are reasonably stable over the past 25 years. Then we can attribute all of the recent high prices in wheat to a declining U.S. dollar. Over the long term, using wheat prices, adjusted by either inflating or deflating the cash, should produce a normal historic probability distribution skewed to the right to show temporary periods of high price.

Structural Changes

Despite the need to correct for inflation, structural changes will affect the smooth pattern of the frequency distribution. For example, the creation of the Eurozone altered the way in which European currencies traded. Trying to analyze the euro using the Deutsche mark prices prior to issuing the euro is comparing two structurally different periods. Similarly, the shift from *live hogs* to *lean hogs* changes the nature of the data.

What about markets that have had major changes in the way investors use them, such as gold and crude oil? Is the gold market that same at $1700/oz as it was at $250/oz when it started trading freely? Is oil the same at $100/bbl as it was when OPEC had a target price of $18/bbl in the 1980s? Technically, there has been no structural change that can be assigned a date and a reason, but the distribution of prices would show that those markets are trading near the far right. If we remove the effect of the U.S. dollar, prices drop by 50% but still trade at the far right of the tail. If we believe that the shape of the normal curve, skewed to the right, is correct, and that prices peak and then fall back to normal, then we must find other ways to correct prices than just removing the change

in the U.S. dollar. For oil, the factors seem to be a *risk premium* due to instability in the Middle East as well as increased demand from Asia due to a rapidly growing middle class. But how do you put values on that?

For a trader, it might be best to simply understand that a return distribution that does not satisfy the normal shape means that prices are in an unusual period and we cannot apply the usual solutions. Attempting to remove demand factors is beyond the scope of this book and could be beyond the scope of many economists. However, these abnormal distributions, when viewed in the short term, present both a trading opportunity and a risk.

Medians and Means

Although the inflation or dollar-adjusted long-term distributions show that prices spend much more time at lower levels, there is a tendency to consider the average price as the midpoint between the highest and lowest prices. But the average gets distorted when there is a run of extreme prices. The median is always a better measurement of the normal price. This was discussed in Chapter 2. For the case of wheat, the average of all monthly prices was $362.75 while the median was $341.50, showing that the recent run-up affects the average much more than the center price. For the wheat series adjusted by the dollar index, the average was 0.00406 and the median 0.00243, also skewed but by not nearly as much. While the median always comes at sorted position 151 for the 302 months, the average corresponded to position 186 for cash wheat and position 157 for the series adjusted by the dollar index (DX). The difference between the locations of the average and median values is a good measurement of the distribution skew. In Chapter 2 the relationship was given as

$$(Skewness)\, S_K = \frac{3 \times Mean - Median}{Standard\ deviation}$$

For price distributions, the median is a much more useful value than the average, although not always as convenient to calculate. The median naturally adjusts for the skewness in the price patterns.

Short-Term Distributions

Unlike the normally predictable patterns of the long-term price distributions, the distribution patterns of short intervals can vary widely; they have been interpreted in many ways. Short-term distributions are not anchored to a base level because the entire period of analysis may be at price levels that are significantly above intrinsic value or production costs. Keeping in mind the normal shape of price distribution for each market group, a different configuration can be a strong indication of expected price change.

Three distribution patterns are shown in Figure 18.9 (b–d) along with the larger, long-term distribution (a). The normal distribution (a) shows the typical clustering at low prices with a long tail toward higher prices. The first short-term distribution (b) mimics the normal pattern but with lower frequency due to less data and the extended tail much
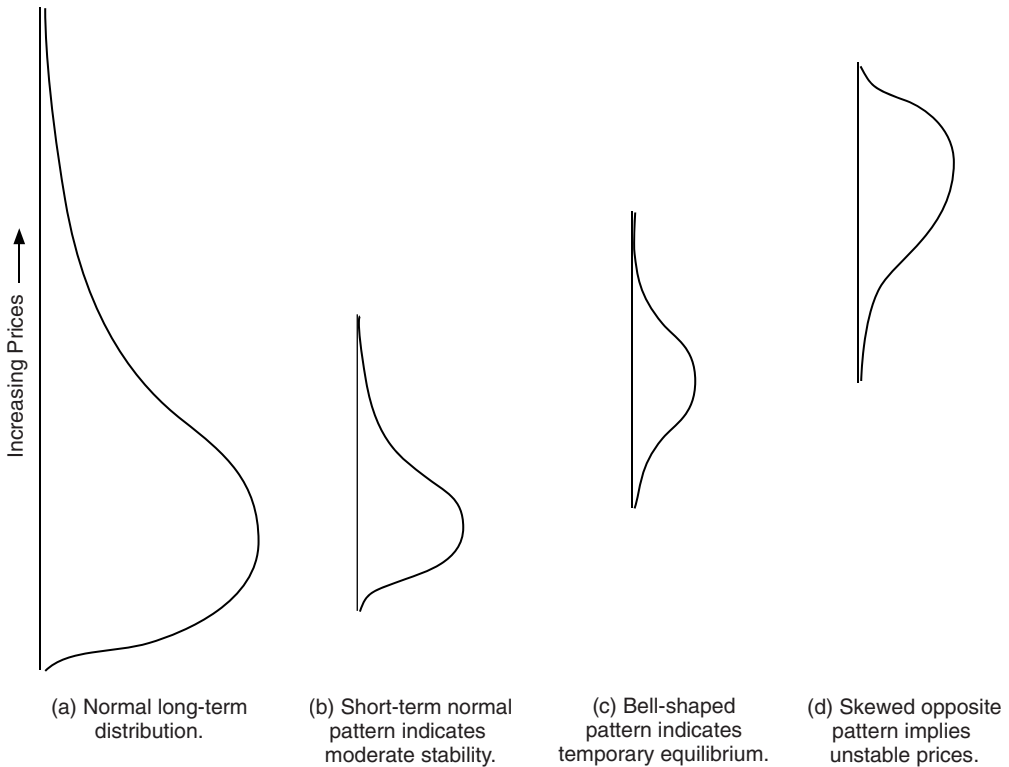
FIGURE 18.9 Three short-term distribution patterns. (a) A normal, long-term distribution. (b) A short-term normal distribution indicating stable prices at a temporarily higher level, or a permanent change. (c) A bell-shaped curve representing congestion but not indicating future direction; the market is in suspension or temporary equilibrium. (d) The tail points to lower prices with most activity at higher prices; the market is unstable and has substantial downside potential.

shorter. This is most likely to occur at low price levels, where most activity is near the bottom with occasional short excursions up. Distribution (b) can be a small segment of a normal pattern or a temporarily higher level that has relatively quiet, stable price movement. Figure 18.9c is a bell-shaped, symmetric curve that could indicate a congestion area or short-term equilibrium at a price level not near the lows but also not at exceptionally high prices. When prices trade equally within a range, they do not exhibit any likelihood of favoring an immediate breakout either up or down; however, in the big picture, declining prices are always more likely.

The fourth distribution (see Figure 18.9d) is most interesting because it is skewed in the direction opposite to (a) and (b). For this pattern to have developed, it required a relatively fast move to a higher price level, then a congestion area formed at the higher level. If you had seen this distribution without knowing the actual prices, you could be certain from this pattern that prices are higher than normal. The pattern is top-heavy, and should be interpreted as unstable; prices are more likely to go down than up. This last

conclusion does not require statistics, because prices at historically high levels may not decline immediately, but must eventually correct to normal levels. Buying into a market with this distribution assures high risk.

Identifying Potential Price Moves

From the patterns just described, a distribution that combines a long tail toward higher prices, with prices that are relatively low, offers no promise of a potential upwards price move. When prices are closer to the long-term *median* or *mode*, the most common occurrence, they are less likely to move. Potential volatility increases as prices get further from the mode. Three distributions that have greater chances for price moves are:

1. A broad, often erratic formation indicates volatile current movement. Prices have not remained at one level, and have not yet established a well-defined pattern or trading range.
2. A skewed distribution (with the tail toward higher prices), the mode near the base level, with the current price at the high end of the tail indicates that prices are in the process of making an unusual, highly volatile move to the upside.
3. A skewed distribution (with the tail toward lower prices), the mode in the upper half of the distribution, and where the current price is in the tail (lower prices) indicates that prices have peaked after a significant move up and are starting down.

Any time prices form a distribution that is not the same as a long-term distribution and prices are not at low levels, there is potential for movement and volatility.

Using Skewness to Identify a Trend

It can be argued that the skewness of closing prices, rather than intraday prices, is a robust measurement because the closing price is the final accounting for the day.[4] Skewness can be used as a leading indicator of a trend change if it is calculated as a *moving skewness* and acted on when it shows early signs of change.

This calculation of moving skewness involves four steps:

1. *Double smoothing* of the closing prices, to get the *moving mean* (M), with $a = 0.1$,

$$S_t = a \times close_t + (1 - a) \times S_{t-1}$$
$$D_t = a \times S_t + (1 - a) \times D_{t-1}$$
$$M_t = ((2 - a) \times S_t - D_t) / (1 - a)$$

---

[4] Dennis McNicholl, "Old Statistical Methods for New Tricks in Analysis," *Futures* (April 2002).

2. *Moving deviation* of the difference between the closing prices and the smoothed mean,

$$DEV_t = Close_t - M_t$$

3. Moving standard deviation of the moving deviation, with $b = 0.05$,

$$SMD_t = b \times |DEV|_t + (1-b) \times SMD_{t-1}$$
$$DMD_t = b \times SMD_t + (1-b) \times DMD_{t-1}$$
$$MD_t = ((2-b) \times SMD_t - DMD_t)/(1-b)$$
$$STD_t = 1.25 \times MD_t$$

4. Moving skewness, with $c = 0.333$,

$$G_t = (c \times DEV_t^3 + (1-c) \times G_{t-1})/STD_t^3$$

The resulting value of G is shown in the lower part of Figure 18.10 applied to the *e*mini S&P. The largest spike in the moving skewness, in July 2009, comes right at the beginning of a new uptrend. The second largest spike, in June 2010, also points to a change of trend, although a bit late. The more frequent smaller spikes are also to be interpreted as turning points, although not as important. The program that calculates the moving skewness is *TSM Moving Skewness*, available on the Companion Website.

Trading Using Kurtosis and Skew

Kurtosis, discussed in Chapter 2, gives the peakedness of the distribution. Because it is the 4th moment, using the power of 4, values are all positive, while the skew can be
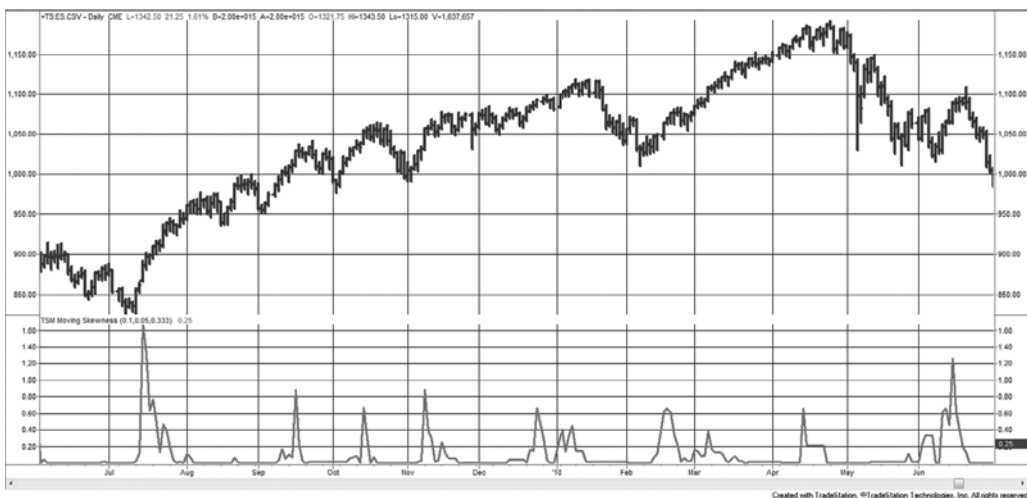
FIGURE 18.10   Moving skewness, *e*mini S&P futures, June 2009 through June 2010.
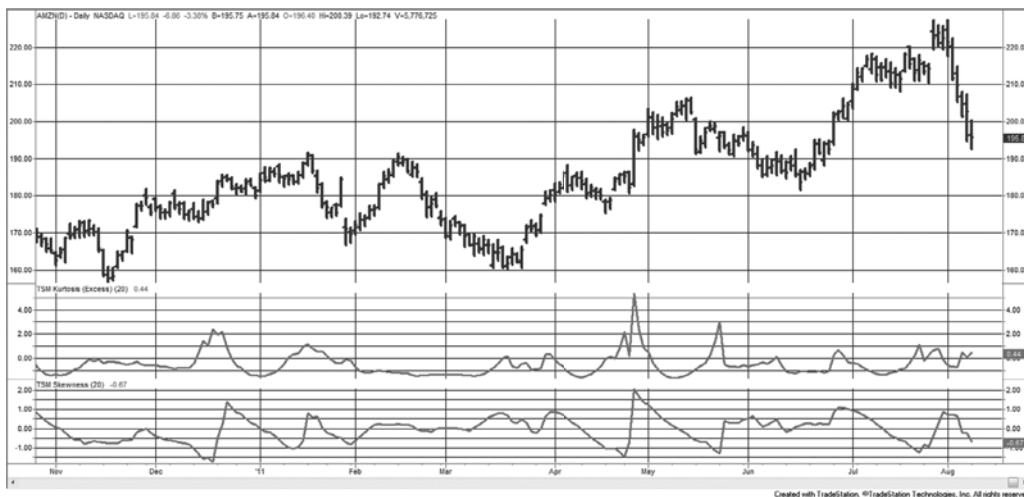
FIGURE 18.11   Excess kurtosis and skew in the bottom panels, applied to Amazon.com during most of 2011.

negative. It is common to show *excess kurtosis*, which is *kurtosis* − 3, where 3 is considered normal. Figure 18.11 shows both moments applied to Amazon (AMZN) during most of 2010.

Note that both the excess kurtosis and skew peak in March 2011, but then show opposite extremes in May and again at the beginning of August. When they are the same upward direction, kurtosis indicates that prices are clustering near the mode, also interpreted as persistence, and the skew shows that the tail is to the right (higher prices). When they peak in the opposite direction, the skew is to the left, toward lower prices. Then two significant peaks will show a likely turning point. Programs to plot these moments are *TSM Kurtosis* and *TSM Skewness*, both available on the Companion Website. If we use kurtosis in a trading strategy, we would want to buy or sell in the direction of the trend when the kurtosis is decreasing, indicating an orderly price move, and we would want to fade (enter in the opposite direction) when the kurtosis is increasing.[5] Skew can also indicate a developing price pattern. If the skew is toward higher prices, and we are trend following, we would buy; if it leans toward lower prices, we sell. Volatility, measured as the average true range, can also help. Lower volatility is associated with lack of movement, while high volatility is often accompanied by large, fast price moves. The rules for a trend-following system that combines kurtosis, skew, and volatility are:

- *Buy* when excess kurtosis crosses below 0, skew > 0, and volatility > minimum level.
- *Sell* when excess kurtosis crosses below 0, skew < 0, and volatility > minimum level.

---

[5] The original TradeStation program for trading based on kurtosis and skew was provided as part of the Omega Research (now TradeStation) System Trading and Development Club. The program in this section is a modified version of that strategy.
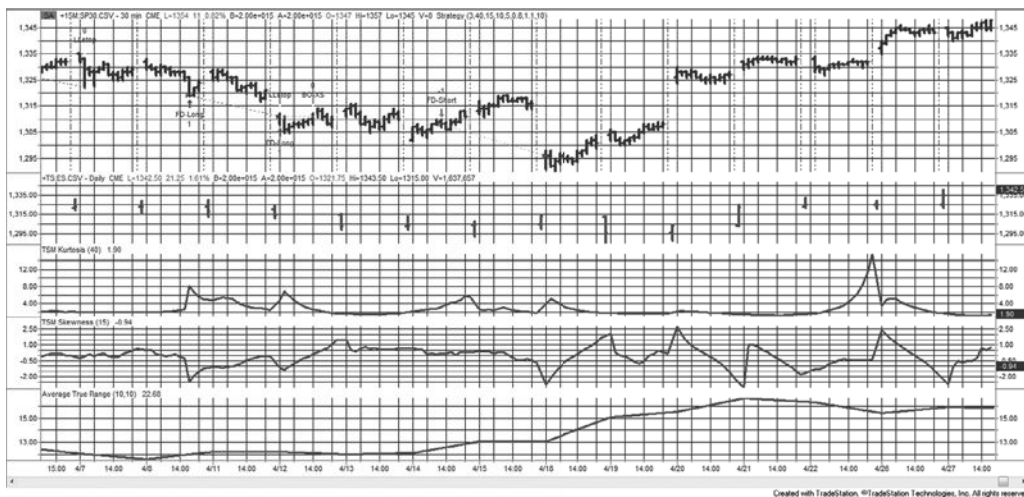
FIGURE 18.12 The Kurtosis-Skew strategy applied to *e*mini S&P, 30-min data (top panel) and daily data (second panel), from March 17 to April 6, 2011. The kurtosis, skew, and average true range are in the bottom 3 panels.

A mean-reverting (fading) strategy can be created by trading when the excess kurtosis crosses above zero, indicating an excessive move, then trading the opposite way indicated by the skew. Volatility will also be important. The rules for mean reverting are:

- *Buy* when the excess kurtosis crosses above 0, skew < 0, and volatility > minimum level.
- *Sell* when excess kurtosis crosses below 0, skew > 0, and volatility > minimum level.

Figure 18.12 shows the Kurtosis-Skew strategy applied to the *e*mini S&P during March and April 2011. Because the kurtosis and skew are calculated over different time periods, the logic from the previous strategy cannot be applied. In this case, the 30-minute S&P is in the top panel, the daily S&P in the second panel, and the 40-bar kurtosis, 15-bar skew, and the 10-day ATR are in the bottom 3 panels. While the volatility is used to decide the trading signal, stop losses are

- (*For longs*) 5 points below the lowest low of the past 10 bars
- (*For shorts*) 5 points above the highest high of the past 10 bars

The 5 points originally used in the stop order for this strategy need to be changed for each market. Because the use of the highest high and lowest low is self-adjusting, a small point value is all that is needed to assure a penetration. It may be worth looking at this with all the calculations spanning the same period and for only daily data. As the frequency of data increases, so does the noise associated with those price moves. The noise may contribute to a less reliable probability distribution, made even less stable by the short calculation periods. Daily data may offset some of these problems.

Additional Strategy Features

The rules for the Kurtosis-Skew strategy used only three variables: the period for kurtosis, the period for skew, and the minimum volatility. There are other features that should be considered:

- A *stop-loss* for long positions based on the lowest low and for short positions based on the highest high of a recent period.
- A *maximum volatility*. Because mean-reverting trades performed better with higher volatility, trend trades may net better profits if they do not enter when volatility is high.
- *Exiting when kurtosis moves the wrong way.* The existing strategy will enter when kurtosis is moving the right way but not exit when it moves the wrong way. Results may be improved by exiting when the excess kurtosis crosses zero, or it may be even better to track the trend of kurtosis, rather than only its relationship to zero.

STEIDLMAYER'S MARKET PROFILE

Market Profile, the effort of J. Peter Steidlmayer, appears to be a frequency distribution of intraday price movement but uses time, rather than volume, as its key element. Steidlmayer formalized this technique in 1985 while he was a Director of the Chicago Board of Trade. The Board considered this such a unique insight that they copyrighted his work as *Market Profile* and *Liquidity Data Bank*. It was intended to provide detailed information to traders about how trading was facilitated. Market Profile offers much more than a count of how many times a price traded at one level. With proper studying, traders are able to separate participants by time frame and even identify their trading patterns. Figure 18.13 shows this technique as it originally appeared, with letters denoting sequential 30-minute time periods, and an analysis of who participated in the trade at each price.[6] Called a *customer trade indicator* (CTI), these categories are CTI 1, the local floor traders; CTI 2, the commercial clearing members; CTI 3, clearing members who fill orders for other members and nonclearing commercial traders; and CTI 4, clearing members who fill orders for the public or other customers not included in the previous groups (*outside paper*). The information on whose behalf the trade was executed is only available after the close of trading.

In order to understand what is described in Market Profile and the Liquidity Data Bank, you must know that the locals and commercial clearing members, CTI 1 and CTI 2, account for the largest part of volume (over 65%), and both have trading styles that are very different from the off-floor speculators who form groups 3 and 4. The locals take smaller positions, often counter to the current price moves and hold that trade for a period spanning seconds to a few hours. Most locals even up at the end of the day, which avoids the need to finance their positions. A *commercial* might be a bank that is hedging

---

[6] F. M. "Doc" Haynie, "Stretching the 'Profile' to Cover 24-Hour Markets," *Futures* (February 1992).

Market Profile
Analysis of a Trading Day

Price	Volume	% Volume by Type of Trader				Cnt	Participation			% Count		% Volume		Ratio (1+2)/(3+4)	
		CTI 1	CTI 2	CTI 3	CTI 4		CTI 1+2	CTI 3+4		%1+2	%3+4	%1+2	%3+4		
9609	736	54.6	8.8	0.0	36.5	2	63.4	36.5	LM	1.3	0.7	467	269	1.74	Buy 3+4
9608	8009	56.2	14.3	7.3	22.3	2	70.5	29.6	LM	1.4	0.6	5646	2371	2.38	
9607	7410	56.9	16.0	8.1	19.0	3	72.9	27.1	KLM	2.2	0.8	5402	2008	2.69	
9606	18142	62.1	13.0	3.5	21.4	2	75.1	24.9	KL	1.5	0.5	13625	4517	3.02	Sell 1+2
9605	31686	62.5	11.7	4.0	21.9	2	74.2	25.9	KL	1.5	0.5	23511	8207	2.86	
9604	13018	58.5	17.4	2.0	21.8	2	75.9	23.8	KL	1.5	0.5	9881	3098	3.19	
9603	14290	60.7	11.7	4.5	23.1	2	72.4	27.6	KL	1.4	0.6	10346	3944	2.62	
9602	5302	52.8	11.8	1.6	33.8	1	64.6	35.4	K	0.6	0.4	3425	1877	1.82	Buy 3+4
9601	3616	69.9	8.7	6.2	15.3	2	78.6	21.5	JK	1.6	0.4	2842	777	3.66	Sell 1+2
9600	4884	50.5	8.4	1.9	39.1	2	58.9	41	JK	1.2	0.8	2877	2002	1.44	Buy 3+4
9531	11828	60.4	12.5	1.2	25.8	1	72.9	27	J	0.7	0.3	8623	3194	2.70	
9530	13464	57.1	16.7	2.5	23.8	1	73.8	26.3	J	0.7	0.3	9936	3541	2.81	
9529	15878	59.7	12.7	3.5	24.1	1	72.4	27.6	J	0.7	0.3	11496	4382	2.62	
9528	4802	57.9	18.1	2.6	21.5	1	76	24.1	J	0.8	0.2	3650	1157	3.15	Buy 1+2
9527	4292	50.4	12.1	1.6	35.9	3	62.5	37.5	GIJ	1.9	1.1	2683	1610	1.67	Buy 3+4
9526	23594	63.1	11.0	4.8	21.1	5	74.1	25.9	FGHIJ	3.7	1.3	17483	6111	2.86	
9525	27090	58.5	15.6	2.7	23.2	5	74.1	25.9	FGHIJ	3.7	1.3	20074	7016	2.86	
9524	20004	59.7	15.3	2.5	22.4	7	75	24.9	CEFGHIJ	5.3	1.7	15003	4981	3.01	
9523	13956	60.1	18.2	3.0	18.7	7	78.3	21.7	BCDEFGH	5.5	1.5	10928	3028	3.61	Sell 1+2
9522	11662	59.4	16.2	5.0	19.3	6	75.6	24.3	ABCDEF	4.5	1.5	8816	2834	3.11	
9521	23390	58.5	14.8	4.5	22.2	7	73.3	26.7	$ABCDEF	5.1	1.9	17145	6245	2.75	
9520	46184	63.5	14.4	5.1	16.9	5	77.9	22	$ABCD	3.9	1.1	35977	10160	3.54	
9519	16018	72.0	9.6	3.7	14.7	4	81.6	18.4	$ABC	3.3	0.7	13071	2947	4.43	Buy 1+2
9518	8750	59.4	12.8	2.3	25.5	2	72.2	27.8	$A	1.4	0.6	6318	2433	2.60	
9517	5912	64.6	12.3	3.9	19.2	3	76.9	23.1	Z$A	2.3	0.7	4546	1366	3.33	
9516	13896	62.9	15.7	4.9	16.5	2	78.6	21.4	Z$	1.6	0.4	10922	2974	3.67	Buy 1+2
9515	5234	62.0	10.1	1.8	26.1	2	72.1	27.9	Z$	1.4	0.6	3774	1460	2.58	
9514	1745	52.9	12.7	0.2	34.3	1	65.6	34.5	$	0.7	0.3	1145	602	1.90	Sell 3+4

Total 374792   Total volume 9528–9518   199742
15%   56219   15% of total
Value area (70%) from 9525 to 9520 (inclusive)   29961.3

R34-29   51555
R7-10   34297
R7-11   65983

FIGURE 18.13   Original Market Profile for 30-year Treasury bonds.

an existing currency exposure or arbitraging a gap between two or more short-term rates in the cash market. These positions may move the market but can be insensitive to current price direction and flow into the market at arbitrary times.

Outside paper, a term that refers to customer accounts, ranges from individual speculators to major fund managers. Because they trade from off the floor, their style of trading favors holding positions for more than one day, and a large majority will be trend followers. The style of CTI 3 is similar to CTI 4, that is, both tend to be directional, and the two can be combined for the purposes of evaluation. Although they account for a smaller portion of the volume, they are more influential in the short-term direction. Figure 18.13 combines CTI 1 with CTI 2 and CTI 3 with CTI 4 to simplify the patterns attributed to each group. During the period shown in the figure, the combined categories 3 and 4 caused the market to move when their participation was greater than 30%. Observation shows that these trading points occur as prices move out of a previous area of sustained trading.

Steidlmayer was said to have first developed this information for his own trading. He taught a course on how to apply his evaluation technique, and there were a number of books written on the subject by himself and others.[7] He spent considerable time studying and classifying formations, and separating days into three primary categories: *normal*, *trending*, and *non-trending*. A normal day forms a standard bell-shaped curve with the widest point positioned near the center; this point is called the *value area*. On a trending day the value area is not as wide and appears closer to one end of the distribution. A non-trending day has neither of the two recognizable patterns.

There is an important philosophic basis for Steidlmayer's work. It is a very credible attempt to explain how the market functions. He says that "the market probes high prices to attract sellers and low prices to attract buyers." This creates a *value area*. By observing the patterns of intraday price distribution, you can know whether there are countertrend opportunities, when prices reach the extremes of perceived value on a normal day, or when trend opportunities exist in the direction of current price movement.

Construction of Market Profile

The *Market Profile* is created in a manner very similar to a point-and-figure chart. You can list every price on the left scale or create boxes that represent a price range. This latter technique will make the distribution look more uniform by clustering the price activity in exactly the same way as a frequency distribution.

---

[7] Three books on Market Profile appeared in 1989, J. Peter Steidlmayer, *Steidlmayer on Markets* (New York: John Wiley & Sons, 1989); Kevin Koy, *Markets 101* (Chicago: MLS Publishing, 401 S. LaSalle St., Chicago, IL 60605, 1989); and Donald Jones, *Applications of the Market Profile* (Chicago: CISCO, 32 S. LaSalle St., Chicago, IL 60604, 1989); however, the most complete material comes from the original course taught by Steidlmayer, called the *Market Logic School*. Readers may also want to study the original publication, *CBOT Market Profile* (1984), available from the Chicago Board of Trade, also J. Peter Steidlmayer and Shera Buyer, *Taking the Data Forward* (Chicago: Market Logic Inc., 1986); and J. Peter Steidlmayer and Kevin Koy, *Markets & Market Logic* (Chicago: The Porcupine Press, 1986).

Value Area Using Daily Data	
Price	**Day Traded**
10228	1
10224	12
10220	120
10216	120
10212	120
10208	245790
10204	2456790
10200	245679
10228	2345679
10224	2345679
10220	23589
10216	589
10212	58

36 day counts (adjacent to rows 10212–10220 center block)

FIGURE 18.14 Daily Market Profile.

Source: Jones, Figure 2, "Locating Value with Auction Market Data," *Technical Analysis of Stocks & Commodities* (July 1989). Used with permission.

The most common way of plotting intraday data for *Market Profile* is to begin by placing the letter "A" in the boxes alongside the price at which the market traded. If we use half-hour intervals, called *time/price opportunities* (TPOs), then all prices that were traded during the first half-hour are marked with the letter "A." The letter "B" is used for the second half-hour period, the letter "C" for the third period, and so on. At the end of the day we have a chart that looks similar to Figure 18.13.[8]

Market Profile is also intended to be used over many days, identifying long-term patterns as congestion areas or trending markets. To the extent that it accomplishes this distinction it can be very valuable. Determining whether a market is directional or not is the trader's greatest challenge. In Figure 18.14 the Market Profile is constructed by using numbers instead of letters to represent the range of prices traded over each day. The entire chart covers a period of 9 days. Other than the substitution of days for half-hour periods, the chart is created in the same way.

Time/Price Opportunities

Time/Price Opportunities (TPOs) are the 30-minute blocks that are assigned letters on the *Market Profile* diagram. It is worth noting that the largest occurrence of TPOs at a single price does not necessarily correspond to the price at which the highest volume occurred, as seen by referring again to Figure 18.13. This distinction is at the root of *Market*

---

[8] Vlarii Salov, "Market Profile and the Distribution of Prices," *Futures* (June 2011), explores the use of tick data and the properties of the distributions.

Profile because it emphasizes the amount of time that traders accepted a price, rather than the volume traded at that price (which could have occurred during a single period). In the bond example given in Figure 18.13, the greatest volume was transacted at 9520 although the center of the value area—where there is the largest record of TPOs—is at 9523, 3 points higher. Clearly, the *Market Profile* analysis is seeking out a different way of observing the actions of traders. In the formation of a value area, if the market moves between two prices, such as 9520 and 9523, because those values bracket the current trading range, each traverse will cause the prices in between to be marked with the same TPO letter, regardless of whether any volume occurred at those prices.

TPOs allow us to judge the potential direction for a move out of the current value area. This can tell a trader to buy strength or sell weakness. The market is said to favor the direction indicated by the upper or lower part of the value area that has the highest TPO count. When performing the TPO count, you must have a well-defined value area and limit the count to within that range. Using Figure 18.13 with the mode at 9523, count the TPOs prior to the interval that begins *J*. There were 16 TPOs above 9523 and 21 below (not including 8$ and 3Z). During time interval *J*, the balance shifted to the upside when it touched 9529, at which point there was a large increase in volume. This could be attributed to CTI 1 and CTI 2 exiting shorts and to a lesser degree, going long. The locals were finally joined by CTI 3 and 4 taking a long position at 9600. An imbalance in the TPO count expresses a willingness for the market to favor the direction given by the largest count, but this applies to a normally distributed value area only.

Typical Patterns

It is generally accepted that the market spends the greatest time in a sideways range, or congestion area, and a relatively small amount of time trending. The sideways period may be defined as an *extended time/price relationship* and the trending period as a *brief time/price relationship*. The extended period creates a value area where traders are willing to buy and sell, and this activity is reflected in higher volume. According to Market Profile,[9]

$$Value = Price \times Time = Volume$$

In this analysis, called *auction theory*, the value area is the place that the market is willing to trade, and is seen in the time value. The value area represents about 70% of the market volume. The center of this time value area is the price at which there are the most TPOs. Because market activity spends at least 80% of the time in this value area, prices tend to rotate about the center. *Rotation* is the term given to price action that moves back and forth above and below this central value, building the pattern of normal distribution similar to a bell-shaped curve.

---

[9] Donald L. Jones, "Locating Value with Auction Market Data," *Technical Analysis of Stocks & Commodities* (July 1989).

What Are the Buyers and Sellers Doing?

The interpretation of the Market Profile is based on the concept that, under normal market conditions, prices rise to attract sellers and fall to attract buyers. There is a very active area of trading at a point called *equilibrium* where commercial buyers and sellers exchange freely because they consider the price *at value* (Figure 18.15a). If there are more buyers than sellers, the price rises to attract additional sellers who feel as though they are getting a price that is above value. However, as prices continue to rise there are fewer buyers because they perceive the market as overpriced.
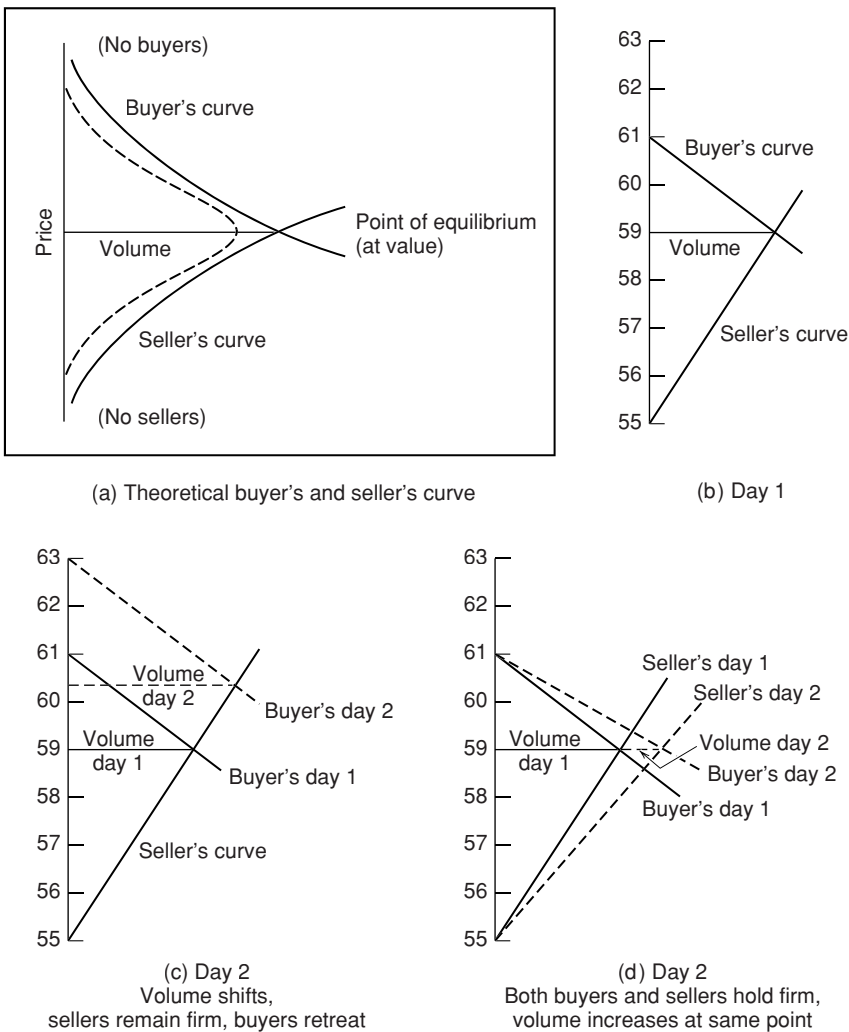
FIGURE 18.15 Buyer's and seller's curves for a sequence of days.

In Figure 18.15a, a buyer's curve and a seller's curve are drawn on the frequency distribution of a normal trading day.[10]

In this way the market is said to *facilitate* trade. When there are not enough buyers, the price falls; when the sellers are scarce, prices rise. Constructing the buyer's and seller's curves for a sequence of days can help understand the dynamics of trading in terms of both price expectations and volume. For convenience, the *curves* are shown as straight lines, similar to supply and demand lines, in the following examples.

In Figure 18.15b, Day 1 begins with the buyer's curve disappearing faster than the seller's; therefore, the buyer's curve is more horizontal and the seller's more vertical. On Day 2 (c), the sellers have held their position, and the buyers are willing to move higher, retreating from their previous objective. This results in higher volume. Had the buyers moved lower and the sellers remained steady, the volume would be expected to decline. If on Day 2 (d) the volume increased while prices traded in the same range and at about the same peak price level (value level), the buyer's and seller's curves would have become more horizontal, indicating that both sides of the trade were holding firmly at the current level.

Quantifying the Value Area

The idea of a value area that is shaped as a bell curve, or an extended formation to represent a trending day, seems clear; unfortunately, actual price patterns are rarely as clear as these examples. To help this process, Jones has applied *overlays* to these patterns, measuring them using a standard deviation. In Chapter 2, we explained that, for a normal distribution, the value of one standard deviation represents a clustering of 68% of all values around the mean; therefore 34% are on each side of the mean. Similarly, two standard deviations contain 95% of all activity, 42.5% on either side of the average. In his analysis, Jones has defined a *value area* as one contained within two standard deviations of the center, using the TPO count to isolate the range. For example, if the entire chart contains 100 filled boxes, the price range is from 6615 to 6655, the center is at 6635, and that row contains 10 filled boxes, then two standard deviations contain $90 \times 0.425 = 38$ boxes on each side of the center. This defines the value area.

Trending Markets

Trends spread volume over a wider price range and are relatively easy to see. In the previous discussions, there is an early warning of a trend when prices move out of the value area or when there are a larger number of TPOs on one side of the center of the value area. Trends develop when the market attempts to probe a new price level, causing the current price to diverge from the value area. The participants may reject this divergence, and price returns to the previous area, causing a broadening formation; alternatively, they

---

[10] The following discussion is based on Robert Pisani, "How Market Structure Helps You Analyze Price Behavior," *Futures* (October 1987).

Beginning of a Trend		
Price	**Day Traded**	
10228	1	
10224	12	
10220	120A	
10216	120A	
10212	120AB	
10208	245790AB	
10204	2456790ABC	
10200	245679ABC	
10228	2345679ABC	
10224	2345679C	
10220	23589C	
10216	589C	
10212	58C	
10208	CD	**Breakout at 10208**
10204	CD	
10200	DEH	
10128	DEGH	**Consolidation or new**
10124	DEGHJ	**value area being formed**
10120	DEGIJ	**around center of 10124**
10116	DEFIJ	
10112	FI	
10108	F	

FIGURE 18.16   Market Profile for a trending market.

may accept the new price as fair value and generate a new value area by attracting volume (see Figure 18.16). The initial trend that fails to find a new value area is called *price trend;* it usually begins and ends within a single time period (one TPO which could be a half-hour or a day). A move that attracts volume occurs over a longer time period and is called a *value trend.* Daily charts, or a sequence of intraday 30-minute charts, are easier to use than a single 30-minute chart when looking for a trend.

When comparing the *Market Profile* of successive days, it is helpful to watch the activity of the recent day compared to the value area of the previous day. When trading within the previous value area, both buyers and sellers are considered equal. A move above the value area is said to be motivated by the buyer and reactive by the seller, and trading below the prior value is initiated by the seller and reactive by the buyer.

Trends do not move in a fast, straight line to their *correct* new level; instead they move in steps, pausing to test whether a new value area can be formed that facilitates price. The trend continues in this way until prices retrace to the previous consolidation area, where the TPO volume is expanded and the area is broadened, resulting in a value area.

FIGURE 18.17   Monkey bars shown on the Thinkorswim trading platform on the top right.

Some Points to Remember about Market Profile

One great advantage of using the TPO count, rather than traditional volume, is that the *Market Profile* can be created at any time on any market. All the necessary information is available. The method is very different from traditional trending approaches and offers insight into the current status of price action, whether it is in a consolidation phase or trending. Although the evaluation of *Market Profile* has been very interpretive, there have been important steps towards defining value areas and trends in a more objective way. Using the overlay method proposed by Jones, a very simple trend filter could be constructed so that new trend trades, based on traditional moving averages, are not set until prices move out of the value area. During the past few years, there seems to have been less interest in this technique; however, its unique concept and sound premise compared with other methods of analysis should keep its standing as a tool worth studying.

Other Trading Platforms

Although not strictly Market Profile, other platforms are able to create intraday price distributions that are similar. In Figure 18.17, *Thinkorswim* plots a distribution they call *monkey bars*. There are three days of 5-minute bars shown for the *e*mini S&P and the bars on the right relate to the individual price distributions for those days.

USING DAILY DISTRIBUTIONS TO IDENTIFY SUPPORT AND RESISTANCE

Support and resistance are most often used to identify breakout points. A new high or low greater than those of the previous *n* days has been a reliable indicator of a trend change. But as prices near those critical levels, is it possible that the price distribution can be a faster and still reliable signal?
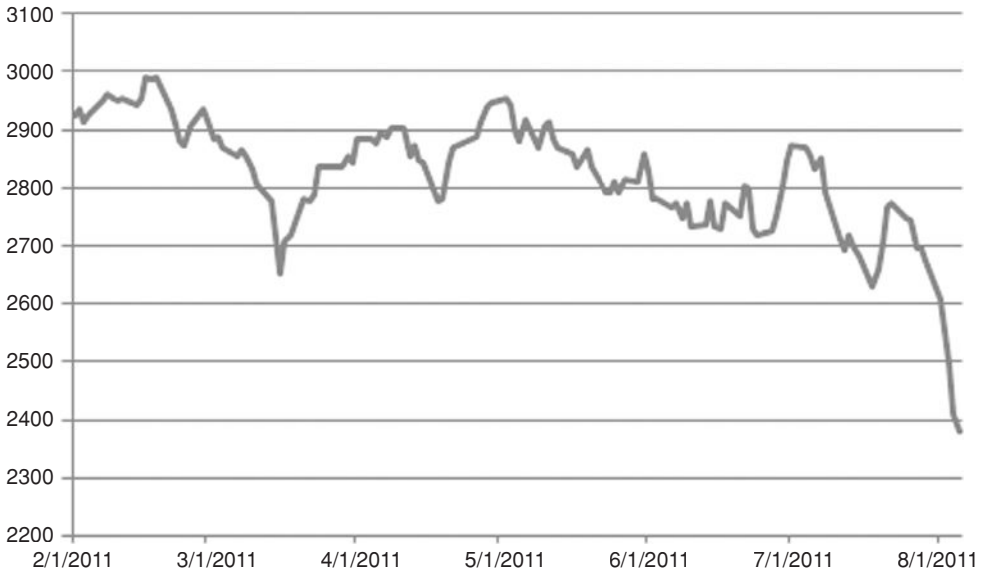
FIGURE 18.18   EuroStoxx futures, closing prices, February–July 2011.

To examine this, the EuroStoxx futures were used from February through July 2011. It was an arbitrary choice. The closing price pattern is shown in Figure 18.18. Most of the chart shows a consolidation area between about 2650 and 3000. Based on traditional chart support, a clear break would occur on a new low below that of mid-March. That occurred in mid-July, followed by a rally and a complete breakdown of prices.

If the same six months are used to plot a frequency distribution, the result is shown in Figure 18.19. To create this histogram, we used Excel's Analysis/Histogram and
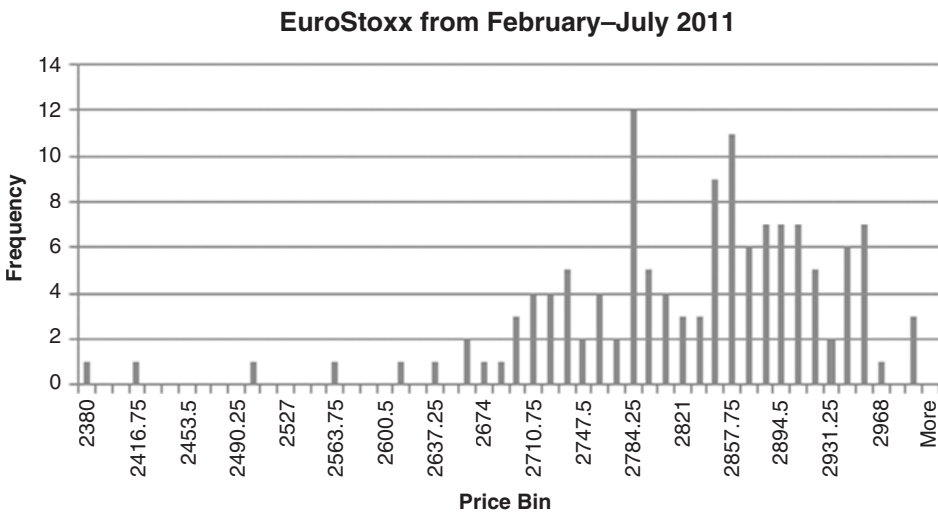
FIGURE 18.19   EuroStoxx frequency distribution of prices, February–July 2011. Vertical bars show expected breakout levels.

created 50 bins from the minimum to the maximum price over the interval, 2380 and 2991. Then each bin was 12.22 points, although we used 12.25. The next step was to define the center cluster. For that we took 60% of the data, leaving 20% on each end. To find those exact points, count the number of prices in the interval (132), then find that price at which the 26th and 106th price fell. That was at 2735 and 2906. Referring to the frequency distribution in Figure 18.19, the 2906 level comes inside the large cluster and does not seem to indicate a breakout level. The minimum, however, at 2735 does seem to show a point where the distribution starts to end and coincides with a series of lows in the price chart. It also shows that prices have made a clear break and are trending lower.

Clearly, this is one case for one market over a short period. However, daily prices also cluster, and a move out of a cluster could be a valid breakout point and a good trading signal. For daily data, longer distribution periods would work well because trading back into a previous range should enhance that pattern. There is no specific percentage that should be used to define the center of a cluster, but 1.5 standard deviations, or about 86% would be a good starting point. If this method gives you a signal ahead of the crowd, then it is valuable.

Multiple Time Frames

I n Chapters 8 and 9, combinations of two trends and two momentum periods were used to create well-known systems and indicators such as Donchian's 5- and 20-day Moving Average, the two-trend crossover systems, double smoothing, the MACD, and the Turtle method, among many others. Each of those methods used daily data for all calculations. This chapter looks at systems where the components use a combination of intraday, daily, and/or weekly data in order to improve both timing and results.

Although the use of multiple time periods by floor traders has been popular for decades, few professionals have talked about it. It is only since better quote equipment has allowed this technique to be accessed by a wider audience that this approach has seeped into the public domain. The combination of multiple time periods allows the trader to time entries into the market using very short-term data, such as 10-minute bars, while watching the longer-term picture for the daily or weekly trend. Because it is agreed that most trends are best identified over a longer time period, and choosing the specific entry point requires a much faster response, the combination of two or even three time intervals is very sensible when each one targets a specific purpose. If the trend can be identified profitably, then the trader can filter or select short-term trades that have a better-than-average chance of becoming winners.

For most traders, the use of any one time frame presents special problems. The very short term contains a high percentage of noise that obscures the market direction. The numerous individual patterns that can be found in a 5-minute bar chart can divert your efforts away from the big picture. The use of only weekly charts, although they clearly show the direction of prices, presents higher risk and little opportunity for a good entry point. The obvious solution is to combine both time frames into a program that uses each to its best advantage.

TUNING TWO TIME FRAMES TO WORK TOGETHER

Throughout most of this book the individual systems and methods have been discussed on their own merits. When using a momentum indicator, an analyst would look for the

specific RSI or stochastic that generated the most profits by showing a timely trend change or overbought/oversold condition. However, very few indicators work on their own as successful trading systems. They are combined with other indicators, and their trading signals are subject to selection using trend lines or divergence.

Once you have decided to combine two techniques, such as a moving average to identify the longer-term trend and an RSI or stochastic to provide improved entry timing, the two techniques must be tuned to work together. The best RSI signals, taken on their own, may no longer be the best entry timing points. You will want the time period for the indicator to be much shorter than the time period used to calculate the trend. For example, if your trend system has a typical holding period of one month (about 23 business days), you will want your timing oscillator to have a good chance of giving you a better entry point within the first 3 to 5 days (10%–20%) of that trade.

To get an indicator to reach overbought and oversold levels an average of every 5 days, you will need to use less than 5 days to calculate the indicator value. It does not matter if the indicator returns a loss when all of its buys and sells are totaled because the profitable part of the program comes from the longer-term trend component. It is most likely that the oscillator will generate five overbought or oversold conditions during the typical holding period for the trend trade, but only one of them will be used—the first one after the new entry signal. It is only important to see how this selective timing helps to improve overall profits and risk of the combined strategy. Readers might remember that Linda Raschke used this technique in *First Cross*.

Figure 19.1 gives an example of the timing needed to coordinate a short-term stochastic indicator with a longer-term trend. The top of the chart shows euro futures (EURUSD) with a 60-day moving average. Trading signals would normally be generated when the
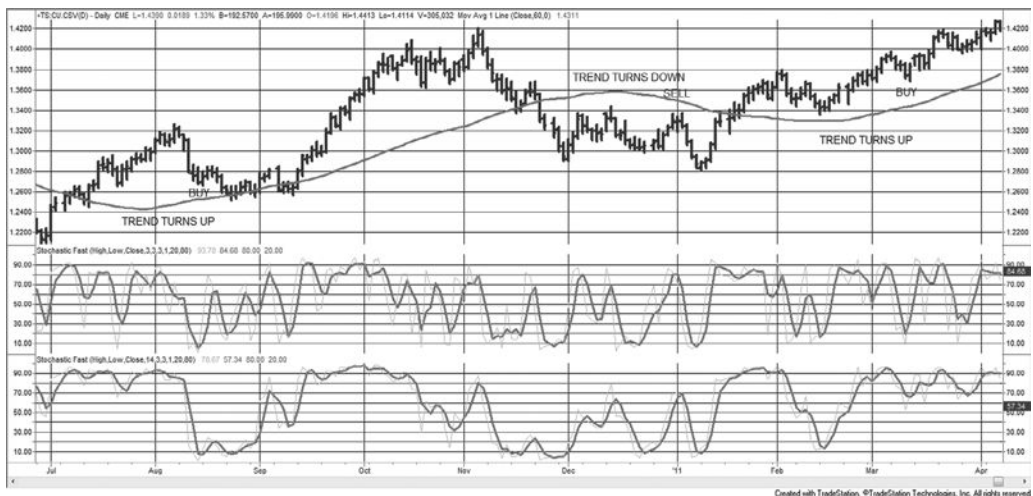
FIGURE 19.1   Timing the entry into a 60-day moving average with a stochastic, applied to euro futures, July 2010 to March 2011. A 3-day raw stochastic is in the center panel and a 14-day raw stochastic at the bottom.

trendline turns up or down. In the first case, the trend turns up on July 30 at a closing price of 1.2986. It turns down on December 16 at 1.3152 for a gain of $2075. Instead, if we used the 3-day raw stochastic in the center panel and waited for the first oversold condition in the stochastic following the upturn, we would have entered on August 11 at 1.2810 and exited on December 31 at 1.3306 for a gain of $6200. At the same time, trades would be entered against the direction of the current price move, generally resulting in less slippage.

Again using the center panel for timing, the long position entered on December 31 would have changed direction on February 22 at 1.3600 using the trend, but on March 10 at 1.3736 waiting for an oversold condition. That causes an entry at a worse price, a situation that happens when prices move steadily in one direction. Overall, entering on a price reversal using a short-term momentum indicator, can greatly improve both performance and the percentage of profitable trades; however, not every entry is better.

Now consider the standard 14-day stochastic in the bottom panel. Had it been used for timing, some trades might have been missed because this longer calculation period has fewer swings from overbought to oversold, and those signals that did occur would have resulted in a significantly delayed entry into the trend trade.

ELDER'S TRIPLE-SCREEN TRADING SYSTEM

The *Triple Screen*[1] popularized a method often used by floor traders who had access to more sophisticated quote and graphics equipment beginning in the early 1980s. It combines trend-following and oscillators using three time frames, each serving a specific purpose. The oscillators are normally associated with timing, while the trend determines the direction of the trade, similar to the process discussed in the previous section. Dr. Elder has observed that each time frame relates to the next by a factor of 5. That is, if you are using daily data as the middle time period, then the shorter interval will be divided into five parts, bars of 1 to 2 hours in length, and the longer period will be five days, or one week.

To be practical, it is not necessary to divide a 6-hour trading day into five intervals of 1 hour and 12 minutes. Rounding to 1 hour is close enough. If, for example, you want to focus on trading a 10-minute chart, then the middle interval is 10 minutes, the short-term is 2 minutes, and the long-term is 1 hour (not 50 minutes).

In the following description, Screen 1 holds the longest time frame while Screen 3 shows the shortest one. Figure 19.2 is an example of the Triple Screen, with 60-minute gold futures in the top panel, daily prices in the second panel, and weekly prices in the third (middle) panel. The bottom two panels show indicators for the intermediate (daily) move, discussed below. In the original Triple Screen, each of the price charts was actually on separate screens, not combined into one.

---

[1] Dr. Alexander Elder, *Trading for a Living* (New York: John Wiley & Sons, 1993).
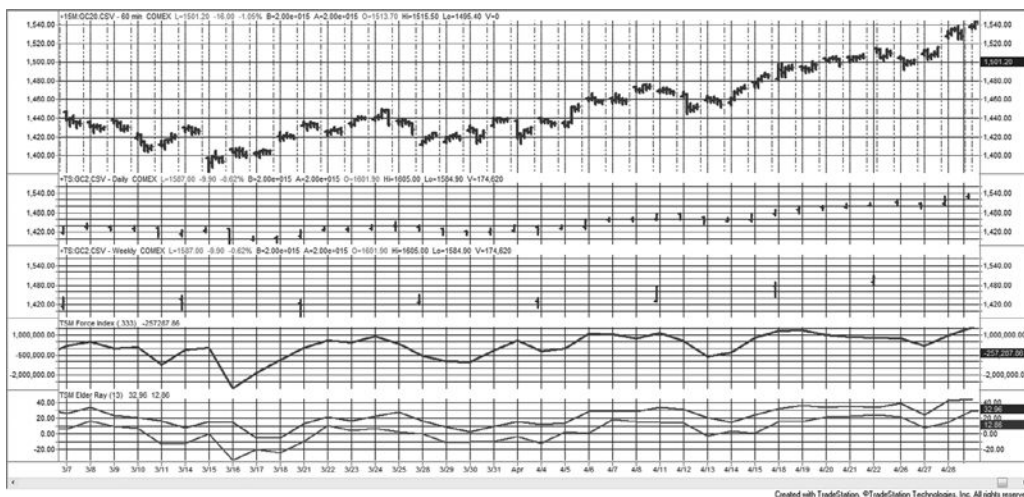
FIGURE 19.2 The Triple Screen for gold futures, March and April 2011. From the top, the panels are 60-minute bars, daily, weekly, the Force Index, and the Elder-Ray.

Screen 1: The Major Move (Lowest Frequency Data)

The long-term view is used to see the *market tide*, a clear perspective of the major market trend, or sometimes the lack of trend. Weekly data is used for this example (the third panel from the top in Figure 19.2), which is consistent with the experience that less frequent data (i.e., weekly or monthly) smoothes the price movement by eliminating interim noise. Although there are many other choices for a long-term trend, the Triple-Screen approach uses the slope of the weekly MACD, where the histogram that represents the MACD value is very smooth, equivalent to, for example, a 13-week exponential smoothing (1 rolling calendar quarter). The trend is up when the MACD bar, or 13-week exponential value, is higher than that of the previous week; the trend is down when this week's value is lower.

Screen 2: The Intermediate Move (Middle Frequency Data)

The oscillator applied in Screen 2 (the second panel) identifies the time period in which we would trade. Again, the specific oscillator is not as important as the time frame and the ability to identify *market waves* in the major moves of Screen 1. Two oscillators are suggested, the *Force Index* and *Elder-Ray*, both described below. A stochastic can also be used.

1. *Force Index*

$$Force\ Index_t = Volume_t \times (Close_t - Close_{t-1})$$

The Force Index is then smoothed using a 2-day exponential smoothing, equivalent to a 0.333 smoothing constant, and the resulting value is used to determine overbought and oversold levels. The Force Index is shown in panel 4.

Entering a long position using the Force Index is not as clear as when using the Elder-Ray. The steps are:

1. Identify the trend in Screen 1, which shows that the major move is up.
2. Confirm the long position if the 2-day exponential of the Force Index falls below its center line and does not fall below the multiweek low. When using a stochastic instead of the Force Index, buy when the stochastic falls below 30.

2. *Elder-Ray*

Elder-Ray is a technique for separating bullish and bearish movement.

$$Bull\ Power_t = High_t - 13\text{-}day\ exponential\ smoothing$$
$$Bear\ Power_t = Low_t - 13\text{-}day\ exponential\ smoothing$$

To determine when to buy using the Elder-Ray and Screen 1, the following two steps are necessary:

1. The trend in Screen 1, the major move, must be up.
2. Bear power is negative but rising; bear power must not be positive.

Two additional steps may be used to filter trades and improve performance, but are not required:

3. The last peak in bull power is higher than the previous peak (the most recent bull power should not be significantly lower than the previous peak).
4. Bear power is rising from a bullish divergence.

The opposite rules apply for sell signals. The Elder-Ray is shown in panel 5 of Figure 19.2.

Both the Force Index and the Elder-Ray can be found on the Companion Website as *TSM Force Index* and *TSM Elder Ray*.

Screen 3: Timing (High Frequency Data)

The final screen is for fastest response, primarily for identifying *intraday breakouts* (see Chapter 16 for a full description of intraday breakouts). It is shown as 60-minute bars in the top panel of Figure 19.2. To improve the point of entry, the intraday bars can be used to set long positions when the current price moves above the previous day's high. There is no calculation involved, simply a *Buy Stop* order using the shortest time period. In this example, a new buy signal occurs when the high of the hourly bar moves above the highest high of the hourly bars of the previous day.

Stop-Loss

Every system needs risk control, and that most often comes as a stop-loss order. The Triple Screen approach positions the stop as a three-step process. For a long position,

1. First place the stop below the low of the day of entry, or the previous day's low, whichever is lower.

2. Move the stop to the break-even level as soon as possible. Naturally, there must be some room between the stop-loss level and the current price; otherwise the stop will always be hit.

3. Continue to change the stop to protect 50% of the highest profits. In addition, you may consider taking profits when the stochastic or Force Index moves above the 70% level.

ROBERT KRAUSZ'S MULTIPLE TIME FRAMES

The most robust and fully automated approach to multiple time frames was taken by Robert Krausz, a well-known trader who provided an extensive course on his method. To understand the importance of first arriving at a sound theory before implementing and testing a trading program, we need to briefly review the characteristics of performance that indicate a robust method.

When testing a trend-following system, we should expect that a trend of 100 days, compared to a trend of 50 days, will produce larger profits per trade, greater reliability, and proportionally fewer trades. As you increase the calculation period, this pattern continues; when you reduce the calculation period this pattern reverses. You are prevented from using very short calculation intervals because relative market noise increases, and slippage and commissions become proportionally larger; on the other hand, the longest periods are undesirable because of large equity swings. When viewing test results as a flowing picture, there must be a clear, profitable pattern and continuity when plotting returns per trade versus the average holding period.

The sophistication in Krausz's work lies in his understanding of this pattern, and its incorporation into the structure of his program, *The Fibonacci Trader*.[2] Krausz also worked in three time frames. Each time frame has a logical purpose, and they are said to be modeled after Gann's concept that the markets are essentially geometric. The shortest time frame is the one in which you will trade; in addition, there are two longer time frames to put each one into proper perspective. The patterns common to time frames are easily compared to fractals; within each time frame is another time frame with very

---

[2]Robert Krausz, *W. D. Gann Treasure Discovered* (Geometric Traders Institute, Fibonacci Trader Corp., 450-106 State Road 13 North, #206, Jacksonville, FL 32259-3863, FibonacciTrader.com). Also see references in the section, "Fibonacci and Human Behavior," Chapter 14.

similar patterns, reacting in much the same way. You cannot have an hourly chart without a 15-minute chart, because the longer time period is composed of shorter periods; if the geometry holds, then characteristics that work in one time frame, such as support and resistance, should work in shorter and longer time frames. Within each time frame, there are unique levels of support and resistance; when they converge across time frames, then the chance of success is increased. In Krausz's work, the relationships between price levels and profit targets are interwoven with Fibonacci ratios and the principles of Gann.

One primary advantage of using multiple time frames is that you can see a pattern develop sooner. A trend that appears on a weekly chart could have been seen first on the daily chart. The same logic follows for other chart formations. Similarly, the application of patterns such as support and resistance is the same within each time frame. When a support line appears at about the same price level in hourly, daily, and weekly charts, it gains importance.

Laws of Multiple Time Frames[3]

As a well-known, successful trader,[4] Krausz brought more than just three time frames and some unique strategies to one display screen. He endowed the program with six rules:

1. Every time frame has its own structure.
2. The higher time frames overrule the lower time frames.
3. Prices in the lower time frame structure tend to respect the energy points of the higher time frame structure.
4. The energy points of support/resistance created by the higher time frame's vibration (prices) can be validated by the action of the lower time periods.
5. The trend created by the next time period enables us to define the tradable trend.
6. What appears to be chaos in one time period can be order in another time period.

Using three time frames of about the same ratio to one another (10-minute, 50-minute, and daily) with daily being the longest, Figure 19.3 shows the June 1998 contract of U.S. bonds with a number of techniques applied over multiple time frames. Figure 19.3a uses only daily bars, while Figure 19.3b uses 10-minute bars; both charts are drawn on the same price scale to facilitate comparison.

---

[3] Copyright, Robert Krausz.
[4] See Jack Schwager, *The New Market Wizards* (New York: John Wiley & Sons, 1992).
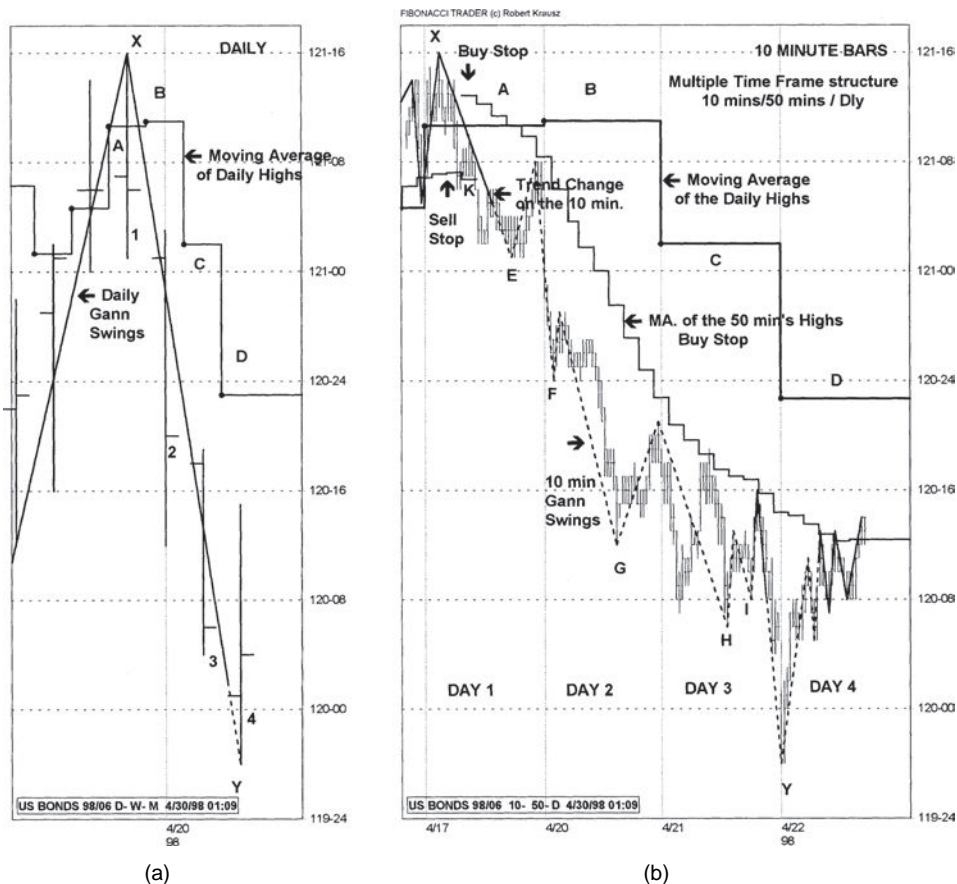
FIGURE 19.3   Krausz's Multiple Time Frames, June 1998 U.S. Bonds. (a) Daily chart with Gann swings and a stepped moving average. (b) Multiple time frame structure for the corresponding four days.

Source: Chart created using *The Fibonacci Trader.* Used with permission from Fibonacci Trader Corporation, 450-106 State Road 13 North, #206, Jacksonville, FL 32259-3863, www. fibonaccitrader.com.

To understand the application of these techniques, it is necessary to identify the following features:[5]

- The *Daily HiLo Activator* (in this case, it is the moving average of the daily highs) is presented as a stepped line. The 4 days of interest are marked by the letters *A*, *B*, *C*, and *D* and appear on both (a) and (b) of Figure 19.3.

---

[5]The following chart analysis was provided by Robert Krausz. Further analysis of bond moves relevant to this discussion can be found in two articles, "A Strategy for Trading Multiple Time Frames," *Futures* (November 2001), and "Intraday Strategies for Multiple Time Frames," *Futures* (January 2002).

- The *50-minute HiLo Activator* (seen in Figure 19.3b) is the moving average of the 50-minute highs, used as a *Buy Stop*, or the moving average of the 50-minute lows used as a *Sell Stop*.
- The *10-minute Gann swings* are based on 10-minute bars (in Figure 19.3b). The solid line shows when the Gann swing represents an upwards trend, and the broken line when it shows a downtrend.

The interpretation of these techniques relies on the faster response provided by the 10-minute bars, combined with the direction given by the longer time frames.

- Based on the 10-minute Gann swings, the trend turned from up to down at about 121-00, while the daily Gann swings placed the trend change much later, near 120-00.
- The slope of the daily Gann swing, measured from point *X* to point *Y* on both charts, was down, defining the dominant trend. Short trades can be entered using the downward trend of the 10-minute time frame. The process of coordinating the trend of the higher time period with that of the lower time period, and acting in only that direction, seems to be the most advantageous approach. The low of each 10-minute swing, marked *E*, *F*, *G*, *H*, and *I* on the 10-minute bar chart, provides opportunities to add to the original position.
- At the top left of the 10-minute chart, the 10-minute close falls below the *Sell Stop* of the 50-minute HiLo Activator at point *K* (about 120-06). The *Buy Stop* then applies and follows declining prices for 3 days. These changes occur in the same area where the Gann swing indicates a trend change from up to down.
- The 50-minute moving average of the highs, shown in a step formation on the 10-minute chart, tracks the highs of the market rallies on Days 1, 3, and 4. The daily moving average of the highs (the Daily HiLo Activator) remained level on Day 2 and turned down on Day 3. The trend can only change to up when the Daily HiLo Activator turns up again.

More on Selecting the Three Time Frames

In thinking out the use of multiple time frames, it is necessary to understand that you cannot substitute a 10-period moving average of 1-hour bars with a 40-period moving average of 15-minute bars. Similarly, you cannot replace a 10-week average with a 50-day average. It seems natural to think that any two trends covering the same time span will give the same result, but that is not the case. Although we can average many data points, we cannot get rid of all the noise; fewer data points over the same time span will always yield a smoother result. Therefore, the use of hourly, daily, and weekly time periods—multiple time frames—gives a much different picture of the market than simply using three different moving averages of the equivalent calculation periods. It is much easier to see the major trend using weekly data, find the short-term direction based on daily data, and time your entry using hourly bars.

MARTIN PRING'S *KST SYSTEM*

Martin Pring approaches multiple time frames using *rate-of-change* (ROC) indicators in his *KSTSystem*.[6] Note that rate-of-change is the same as returns multiplied by 100 (i.e., whole percent). Because each ROC applied to a specific calculation period has a unique cycle, the combination of three calculation periods will provide a valuable confirmation of a trading signal. Although Pring prefers to begin with a long-term view of the market, using 6-, 12-, and 24-*month* ROC calculation periods, Figure 19.4 shows the 6-, 12-, and 24-*week* ROC applied to the S&P 500 futures. Vertical lines show where the cycles of the three time frames align at major and intermediate tops and bottoms. These points will form the basis of the trading rules.

The cycles formed by the ROC are used in conjunction with a trend. The cycles not only confirm the trend change, but the size and frequency of the cycles will help to set objectives. The best trend moves occur when all three ROC indicators are moving in the same direction following a coordinated turning point.

The second stage in the process is to smooth the ROC indicators, making them more reliable. Using the 6-, 12-, and 24-month ROC preferred by Pring, the 6- and 12-month indicators are smoothed by a 6-period moving average, and the 24-month smoothed by a 9-period moving average. Because the ROC calculation actually speeds up the price movement, instead of introducing a lag, this small amount of smoothing shows very little lag
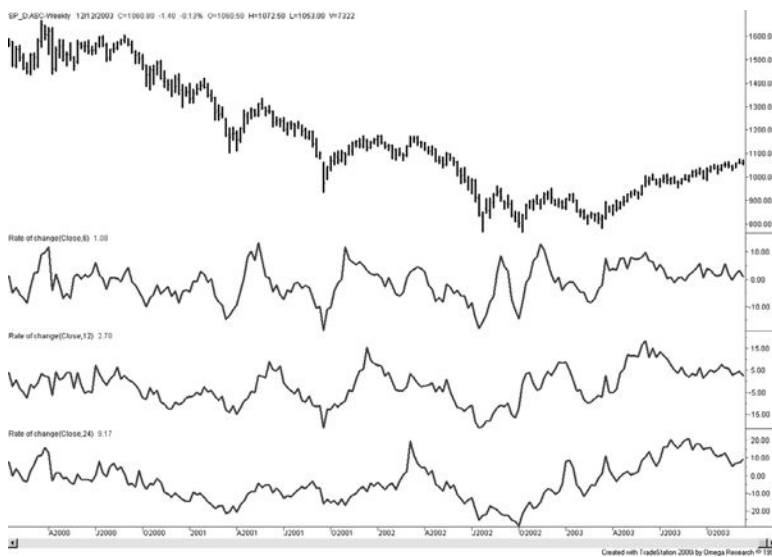
FIGURE 19.4   6-, 12-, and 24-week ROC (top to bottom) applied to the S&P 500 continuous futures, January 2000 to December 2003.

[6]Martin J. Pring, *Martin Pring on Market Momentum* (Gloucester, VA: International Institute for Economic Research, 1993). Also see www.pring.com.

in the smoothed values. Using the three smoothed ROC indicators, Pring selects trades using the following guidelines:

1. A trendline based on about half of the longest ROC period determines the direction of the trade.
2. The longer 24-period ROC identifies the major move.
3. The strongest price moves occur when all three ROC values are moving in the same direction.
4. If the 24-period ROC peaks while the other two indicators are rising, the sell-off is minor. Similarly, when the 12- and 6-period indicators peak but the 24-period is rising, the sell-off is also mild.

    We can conclude that all three indicators need to peak at the same time, followed by all three values declining, to see a significant sell-off.

Creating a Composite Indicator

Pring observed that trading signals generated by combining the ROC indicators with a trend reversal were out of phase with the major moves in the S&P. To correct that problem, he created the *KST indicator*, a composite of four smoothed ROC calculation periods, each step-weighted in proportion to its period.

1. A 24-period ROC weighted 4.
2. An 18-period ROC weighted 3.
3. A 12-period ROC weighted 2.
4. A 9-period ROC weighted 1.

$$KST = (4 \times Average(ROC(price,24),\ 9)) + (3 \times Average(ROC(price,\ 18),\ 9))$$
$$+ (2 \times Average(ROC(price,\ 12),\ 6)) + (1 \times Average(ROC(price,6),\ 6))$$

where $ROC(price,24)$ is a function that returns the 24-period ROC of the *price* series $100 \times \left(\dfrac{p_t}{p_{t-24}} - 1 \right)$, and *Average* (*series*,9) is a function that returns the 9-period average of *series*. In this formula, *series* is the result of the ROC calculation. In his description of KST, Pring does not give the weighting of 18-period ROC; the value 9 is used here.

    To generate trading signals, Pring plots the 9-month ROC along with the KST indicator, plus the 12-period moving average trendline. In Figure 19.5 the bottom panel has an 18-period ROC smoothed using a 9-period average, one of the KST components. Using weekly data, the 18-period ROC seemed to track better. Trading signals are taken in the direction of the moving average trendline in the upper panel, but timed when the KST indicator crosses the ROC line after the first turn in the trendline. With the proper choice of ROC periods, the trade could be exited when the KST crosses the ROC in the other

FIGURE 19.5   KST combined with an 18-period ROC and a 12-period trendline
can improve timing and selection of major price moves. Prices are S&P 500
continuous futures, weekly data, January 2000 to December 2003.

direction. If the ROC is too fast, then it will be difficult to take bigger profits from the
trend moves; therefore, it seems preferable to use one of the slower ROC calculations.

The major trends are clearer, and the signals more reliable, when applied to monthly
data, although the method is generalized to work using any size data bar. A faster version
of KST can be constructed from daily data and used to filter trades or improve their
timing. It is a requirement, however, that the KST signals be used in combination with a
trendline.

The KST is an exceptionally smooth indicator with very little lag, making its signals
more timely than most other indicators. It can also be used in the same way as an RSI
or stochastic, applying divergence and trendlines to generate additional, reliable trading
signals. In Figure 19.5, a trendline has been drawn across the tops of the KST indicator
beginning in April 2002, showing the breakout of a descending pattern in April 2003, a
very timely entry into the stock market.

The KST indicator is available on the Companion Website as *TSM KST*.

Advanced Techniques

After price, volatility has the greatest impact on trading, so it is worth spending time understanding how to measure it and how to use it when trading. Those methods cover a wide range. Volatility has already been discussed as part of many indicators and as a way to make stop-losses, profit-taking, and even the charting scale variable (see the discussion of point-and-figure in Chapter 5) so that the trading methods adjust automatically to changing market conditions.

Volatility can be used as a trading filter, avoiding high risk or standing aside when the volatility is too low. It is the key measure of risk and will be the dominant ingredient in structuring a portfolio, covered in Chapters 23 and 24. As systematic programs mature, there seems to be a greater, justifiable concentration on how to include and manage volatility. Here, we will look at it in more detail.

MEASURING VOLATILITY

In general, the volatility of most price series, whether stocks, financial markets, commodities, or the spread between two series, is directly proportional to the increase and decrease in the price level. Higher prices translate into higher volatility. This *price-volatility relationship* has been described as *lognormal* in the stock market, and is similar to a percentage-growth relationship.

Consider Google in Figure 20.1a, a plot of the price since inception, and 1b, a semi-log plot, available on Excel under *format axis/axis scale/logarithmic*. In a semi-log plot, the *y*-axis is plotted on a log scale, where each equal space up is 10 times the previous level. By showing them side by side, the log scale is clearly less volatile. The faster moves up and down near the highest levels are much smaller on the log chart. Then using a log relationship is one way of explaining the relationship between price and volatility.

FIGURE 20.1 A comparison of (a) Google prices and (b) a semi-log plot of the same prices.

The Price-Volatility Relationship

Before getting immersed in various measurements of volatility, be aware that two stocks trading at the same price can have very different volatility. It may not only be the nature of their business, but those stocks in the news tend to have more volatility. This becomes important in many ways:

- When sizing positions in order to equalize risk.
- When balancing position sizes for pairs trading or market neutral baskets.
- For assessing risk in portfolios.

While each of these issues is covered in other places in this book, it is volatility that is the key driver of risk, and measuring volatility correctly is necessary to control that risk.

Adjusting for a Base Price

While a stock price can go to zero, commodities always have an intrinsic value. Wheat, for example, has a production cost below which farmers are reluctant to sell. During times of surplus, or when an individual farmer is in need of cash, wheat could be sold below cost, but not at zero, and not usually far below cost. When prices get very low because of exceptionally high inventories, they tend to settle near the cost of production and volatility drops significantly. Farmers are not interested in selling so there is little activity. Speculators do not see prices changing until there is more information about the next year's crop; therefore, they also stand aside. From this we can say that, somewhere below the cost of production, volatility goes to zero. Call that level the *base price* for the purpose of calculating volatility.

Then prices can be adjusted by dividing the current price by the base price (p_0) or subtracting the base price from the current price, then taking the natural log (ln) of the adjusted value.

$$Adj \ p'_t = \ln \frac{p_t}{p_0} \quad \text{or} \quad Adj \ p'_t = \ln(p_t - p_0)$$

The charts of either of these adjusted prices will look similar to the unadjusted semi-log plot; however, when the base price is high, the adjusted price will be more useful.

Exceptions

There are three noteworthy exceptions to the price-volatility relationship, and these exceptions will lead to a more general way of measuring volatility.

1. Interest rates trade as prices on the futures markets, which is inverse to the yield. When evaluating long-term volatility for rates, and in most cases when using percentages, yield should be used.

2. Foreign exchange has no base price, only equilibrium, the price that all traders, and the governments, accept as fair value for the moment. This situation is always short-lived. Prices get more volatile when they move away from equilibrium in either direction.

3. Energy is controlled by a cartel. They attempt to set the supply and target a price range. This was very effective when prices were in the range $18 to $32/bbl, but does not seem to have much effect at higher levels. Changing the supply can cause shocks to the market in either direction.

Determining the Base Price

Finding the real base price of a commodity can have uses beyond volatility, in fundamental analysis. There are two straightforward ways this can be done.

1. *Use a linear regression of prices.* Apply a standard least squares regression, available on Excel in Data/Data Analysis, with sequential integers as x and the closing prices
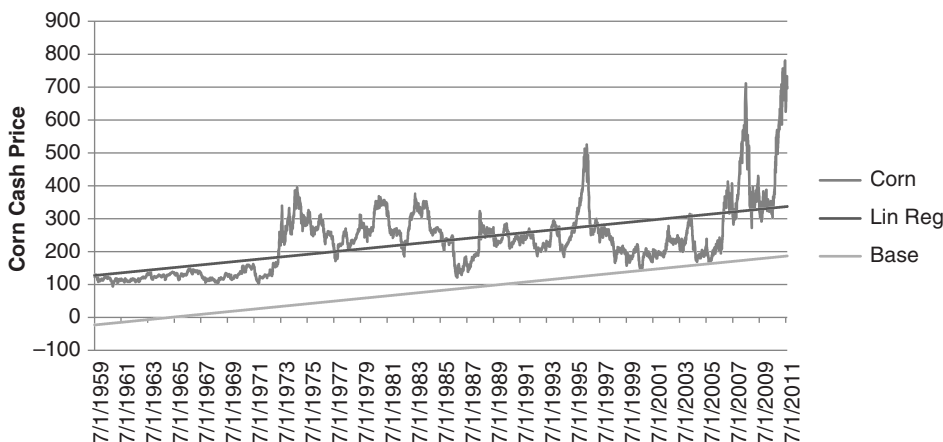
FIGURE 20.2 Using the linear regression to define the base price level.

as y. Then create the regression line as $p_t' = a + bx$, where a is the y-intercept and b the slope. To find the base price, calculate the residuals, $res_t = p_t - p_t'$, then find the minimum residual value. Subtract that minimum value from the regression line to get the base price line, shown in Figure 20.2. The base price continues to increase with time.

2. *Use a linear regression of volatility.* Instead of price, using a measure of volatility related to the price level will give a more direct view of the price-volatility relationship.

Consider monthly copper prices and two measurements of volatility, the returns and the price range for that month. These numbers will not be annualized because that is only multiplying them by a constant, the square root of 12. From a spreadsheet of copper cash prices, calculate the returns of the past 21 days whenever the month changes. At the same time, find the difference between the highest high and the lowest low price for the past 21 days. Put them into a table along with the monthly ending price. Using Excel, choose a scatter diagram with a linear regression line for the combinations of price with returns, and price with range. Both of these are shown in Figures 20.3a and b. Using the monthly price ranges gives a clear picture of the price-volatility relationship, increasing in direct proportion to price and going to zero somewhere near $0.50/lb. Using returns is not as clear, with a large indistinguishable cluster in the bottom left corner and low volatility occurring at high prices. Using the price range whenever possible turns out to be a more robust solution because it has the advantage of using more data, namely the high and lows rather than only the close.

There are many other ways to find the base price and the price-volatility relationship; however, simple is often best. In this case, a linear regression shows the relationship clearly, even if a nonlinear fit may have been more accurate. For the purpose of using volatility in trading systems, the comparison of returns to price range shows
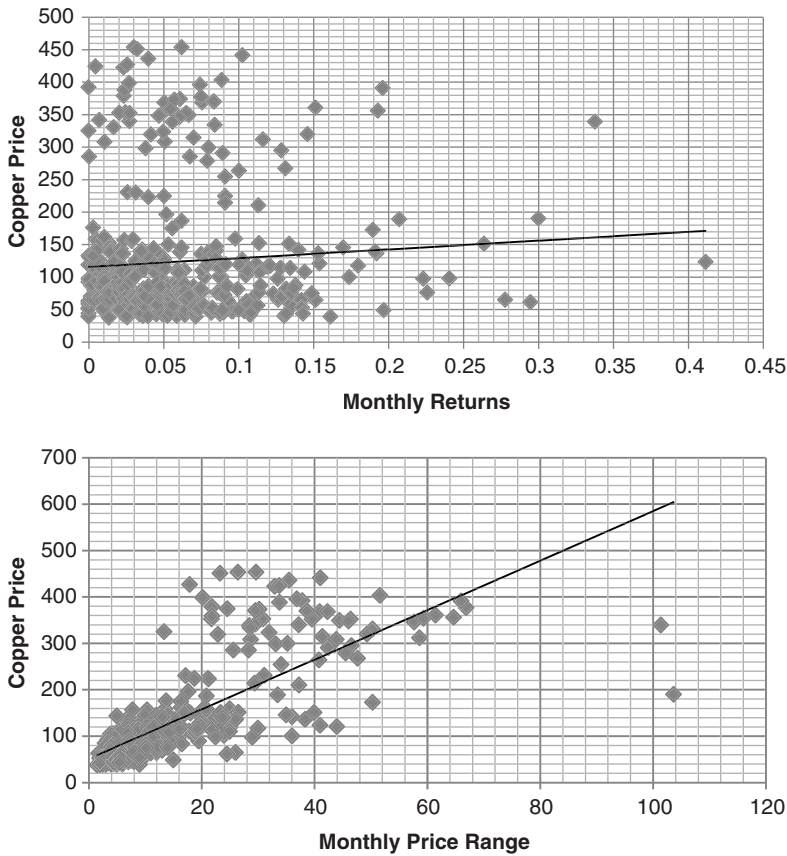
FIGURE 20.3 Scatter diagrams of volatility against price for monthly copper, 1974 to July 2011. (a) Price and monthly returns. (b) Price and monthly price range.

that the price range is a more effective measurement when high and low data are available.

The Time Interval

The time period over which volatility is measured is also a significant factor in the price-volatility relationship. A longer period means that the net changes over n days may turn into months or years. Longer measurement periods give higher volatility values; however, the rate at which volatility increases will decline over time. The maximum time period over which volatility increases will be equal to the period of the largest price move, taken as a percentage. Options traders are familiar with this relationship because option pricing is all about volatility, how much could a stock or futures price move before expiration. Figure 20.4 shows a general representation of how volatility increases over time. This may vary slightly for specific markets.
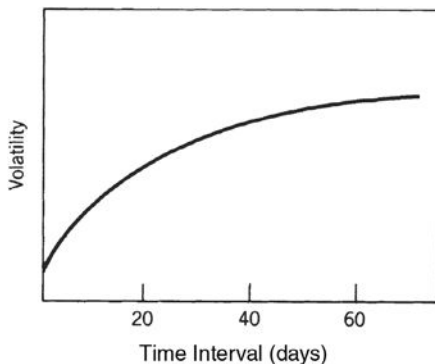
FIGURE 20.4 Change in volatility relative to the interval over which it is measured.

An Example of a Lognormal Calculation

A lognormal relationship in prices can be shown using the math function *natural log, Ln,* in Excel. Over the long term and under average market conditions, the relationship between actual price changes and volatility is expected to be:

$$\frac{|p_t - p_{t-1}|}{\ln p_t} \approx \frac{|p_s - p_{s-1}|}{\ln p_s}$$

This formula says that the absolute value of the change in price on day t divided by the natural log of the price on day t is approximately equal to ($\approx$) the change in price on any other day, s, divided by the natural log of the price on day s.

For example, if the price on day t is 20 and the price on day s is 40, then the natural log of prices on those days are ln (20) = 2.99 and ln (40) = 3.99. If the volatility is $1.00 when $p_t = 20$, then the volatility is expected to be $1.23 when $p_s = 40$.

$$\frac{volatility\ at\ p_t = 20}{\ln(20)} \approx \frac{volatility\ at\ p_s = 40}{\ln(40)}$$

If the volatility at 20 is 1.00, then the volatility at 40 is 1.23. When using a spreadsheet for your calculations, note that the function *ln* is not the same as the function *log*.

Volatility Measures

There are five practical measures of volatility that can be easily used to satisfy the need to show expanding and contracting volatility, but are not tied to an underlying price level. They simply measure volatility over the most recent n days, or n bars using price
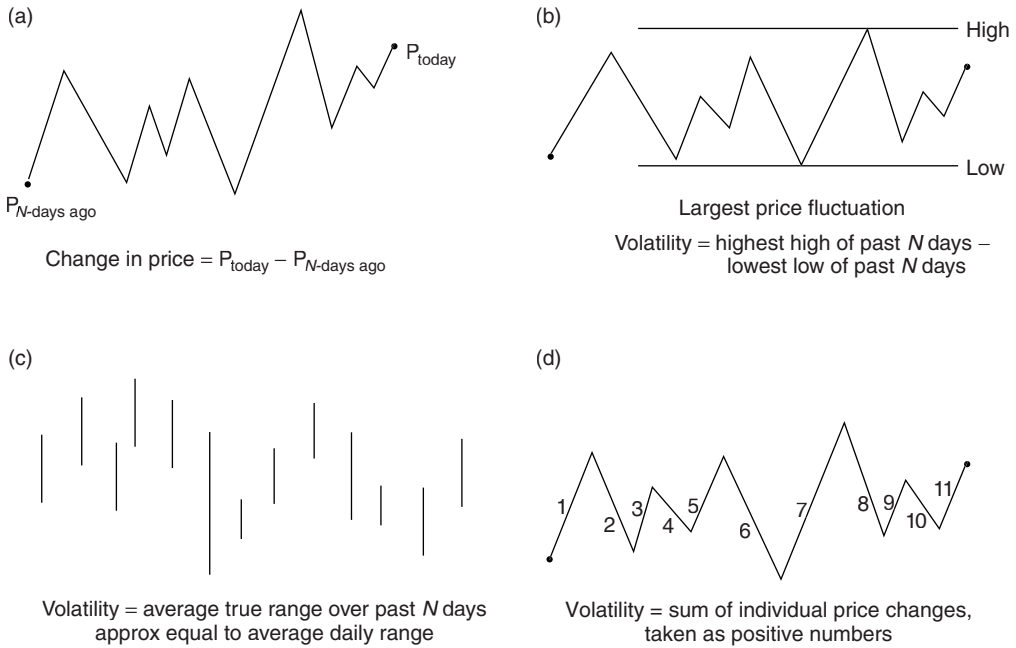
FIGURE 20.5 Four volatility measures.

Source: Perry J. Kaufman, *A Short Course in Technical Trading* (Hoboken, NJ: John Wiley & Sons, 2003). Reprinted with permission.

differences or ranges. Referring to Figure 20.5 in conjunction with the formulas below, the five volatility measures for today, V_t, calculated using the past n days, are:

1. The *change in price* over n days (Figure 20.5a):

$$V_t = Close_t - Close_{t-n}$$

2. The *maximum price fluctuation* during the n days (Figure 20.3b):

$$V_t = Max(High_t, \dots, High_{t-n}) - Min(Low_t, \dots, Low_{t-n})$$

where *Max* and *Min* are the same as the *TradeStation* functions *Highest* and *Lowest*.

3. The *average true range* over the past n days:

$$V_t = Average(Truerange, n)$$

where *Truerange* is a function that returns the maximum range from the combination of today's high, low, and previous close.

4. The *sum of the absolute price changes* over n days:

$$V_t = \sum_{i=t-n+1}^{t} |Close_t - Close_{t-1}|$$

For stocks, the sum of the returns should be used.

5. Classic *annualized volatility* for daily data

$$V_t = Stdev(r_i) \times \sqrt{252}, \text{ where } i = t - n + 1, t$$

In (1), the volatility is entirely dependent on the value of the two points, p_t and p_{t-n}, regardless of the price activity that occurred during the days between them. If prices are very volatile but close to where they began n days ago, this method would show zero volatility. Over many calculations, this method returns a reasonable measure of volatility, but will always be lower than those using the highs and lows.

The maximum range (2) corrects for the dependence on only two points and will produce a more meaningful measure of volatility, which may also be used as an estimate of risk. This method may be effective as the basis for a stop-loss or profit target if you know the average holding period of the trade, because it estimates the maximum move for that period.

The average true range (3) is the most popular measurement of volatility and gives a reasonable guideline for future volatility. Many strategies use a factor times the average true range to place a stop-loss, take profits, or decide the current level of risk.

The sum of the absolute price changes (4) is similar to the average true range but misses the connection to the previous day. Then a large gap that does not retrace to the previous close will appear to have lower volatility. This may be better for intraday trading.

The last method (5), *annualized volatility*, is the most popular and accepted among financial analysts. It is the standard deviation of the returns over the past n periods, times the square root of the number of periods, P, in a year. The generalized formula is

$$V_t = \sigma(r_i) \times \sqrt{P}$$

For daily data based on a 60-day rolling calculation, $i = t - 59$ to t and $P = 252$. For monthly data, $P = 12$. In the case of back-adjusted futures prices, the returns, $r_t = p_t/p_{t-1} - 1$ *or* $\ln(pt/p_{t-1})$ will be incorrect because older data gets distorted due to the continuous adjustment.

Of the five measurements, the two of most interest are the average true range and annualized volatility.

Comparing Annualize Volatility and Average True Range

Results from annualized volatility and average true range can be very different and can significantly affect trading decisions, position size, and risk assessment. Figure 20.6 shows a daily chart of Merck (MRK) from April 2010 to May 2011. The average true range (ATR) is shown in the center panel and the annualized volatility in the lower panel, both
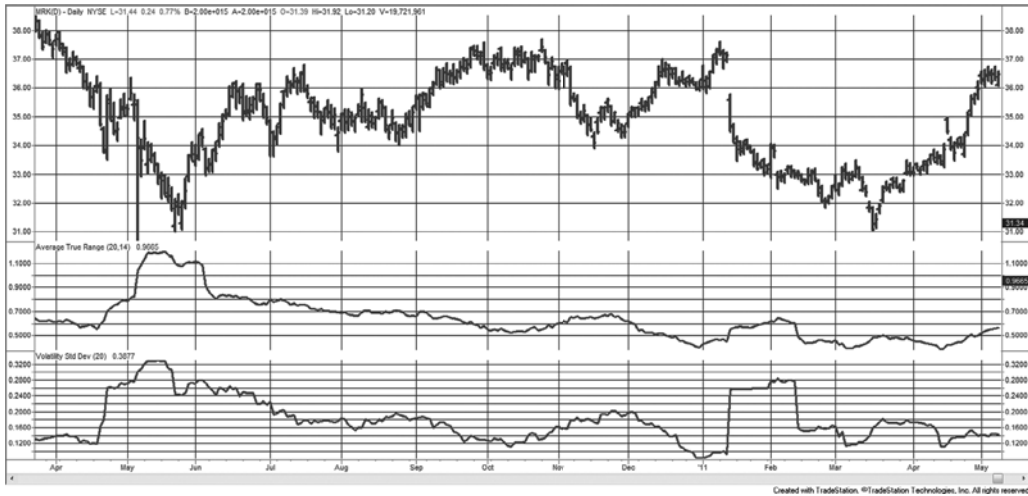
FIGURE 20.6 Merck (MRK) with a 20-day average true range (center panel) and a 20-day annualized volatility (bottom panel).

calculated over 20 days. The ATR is considerably smoother on a daily basis and shows smaller jumps when prices gap. In some cases, such as the far right, the true range is increasing while the annualized volatility is flat. Both can be converted to dollar values by multiplying by the current price.

Ratio Measurements

Bookstaber[1] presents a volatility measurement *V*, in the traditional notation of a return series, that is, the ratio of successive closing prices. He adds alternate measures using the high and low, or a combination of the three as follows,

where C_t = the closing price on day t
  H_t = the high price on day t
  L_t = the low on day t
  V_t = the volatility on day t

(a) Close-to-close volatility (for stock prices or cash data)

$$R_t = \frac{C_t}{C_{t-1}}$$

and

$$A = \frac{1}{n} \sum_{i=1}^{n} \ln R_{t-i}$$

---

[1] Richard Bookstaber, *The Complete Investment Book* (Glenview, IL: Scott, Foresman, 1985), 349.

then

$$V_t^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\ln R_{t-1} - A)^2$$

and

$$V_t = \sqrt{V_t^2}$$

(b) High-low volatility

$$V_t = \frac{0.601}{n} \sum_{i=1}^{n} \ln\left(\frac{C_t}{C_{t-1}}\right)^2$$

(c) High-low-close volatility

$$S_t^2 = 0.5 \ln\left(\frac{H_t}{L_t}\right)^2 - 0.39 \ln\left(\frac{C_t}{C_{t-1}}\right)^2$$

$$V_t^2 = \frac{1}{n} \sum_{n}^{n} S_{t-i}^2$$

$$V_t = \sqrt{V_t^2}$$

Note that t may be a time interval rather than a single day. Then C_t is the last price of the period, and H_t and L_t are the highest and lowest prices of the interval.

In the close-to-close estimation, the volatility V_t is the standard deviation of the closing price ratios. Bookstaber states that this measurement will follow a χ^2 distribution and that the actual volatility during the current period t can be set within the error bounds defined by the distribution.

Relative Volatility

It is often convenient to filter trades by either high or low volatility. Sometimes, the absolute lows and highs are important, but relative highs and lows may work in many cases. *Relative volatility* (*RV*) can be defined as the volatility over a short period divided by the volatility over a longer period, where the longer period is typical of the normal volatility,

$$RV_t = \frac{V_t(n)}{V_t(m)}$$

where V_t is any of the volatility measures, and n and m are calculation periods, where $f \times n = m$, where $f >= 5$, but $f >= 10$ would be better.

Lagging the Longer Period

One problem with the relative volatility, and many other comparative measures, is that the shorter calculation period is contained in the longer one. A better measure would be to lag the longer calculation so that it ends before the shorter one starts,

$$RV_t = \frac{V_t(n),\ i = t - n + 1, t}{V_t(m),\ j = t - m - n + 1, t - n}$$

Then the shorter calculation, n, goes from $t - n + 1$ to t, and the longer one, m, goes from $t - m - n + 1$ to $t - n$, non-overlapping periods. This method will also help on the back side of a volatile period, when the typical calculation includes the recent volatility, making the declining volatility seem normal, rather than still volatile.

Implied Volatility, *VIX*

Even easier than calculating the volatility yourself is to have a good measurement accessible at all times. The CBOE's volatility index, *VIX*, reflecting the *implied volatility* of the S&P options, is available on a real-time basis. The VIX was introduced in 1993 and only recently started trading as futures, valued at 1000 times the price. VIX was originally the volatility of an options index, OEX, a weighted value of the implied volatilities of 8 puts and calls in the S&P 100, expressed as a percentage of the index price. However, the futures contract is constructed from the forward 3-day volatilities of the S&P 500.

Although the OEX is no longer quoted, it is important to understand how the value of the VIX is calculated. If the VIX is 25%, and the SPX is 1600, then VIX is forecasting 25% volatility for at-the-money options, relative to the price of the SPX for 30-day rolling expiration period.[2] Because the 30-day calendar period is equivalent to about 21 trading days, and there are 252 trading days in the year, 1 standard deviation of the volatility (equal to a 25% move in the SPX at 1600) becomes

$$Implied\ range = 1600 \times 25\% \times \sqrt{\frac{21}{252}} = 115.47$$

Then, an implied volatility of 25% when the SPX is at 1600 is equal to a 68% (1 standard deviation) chance of a price change of ±115.47 within the next 30 calendar days (21 trading days).

The VIX is based on a very specific set of criteria that make it generic, rather than specific to any index or stock; however, its relative changes and extreme highs and low can make it very useful for trading. Figure 20.7 shows the VIX plotted below the cash S&P 500 index, SPX, for the now-famous period, March 2008 through March 2009. When the S&P hits its lowest price on October 24, 2008, VIX peaks at 89.53. Previously the VIX

---

[2] Andrew Fisher, "Teaching the Old VIX New Tricks," *Futures* (August 2001). Fisher explains the VIX calculations in addition to proposed improvements in iVIX.
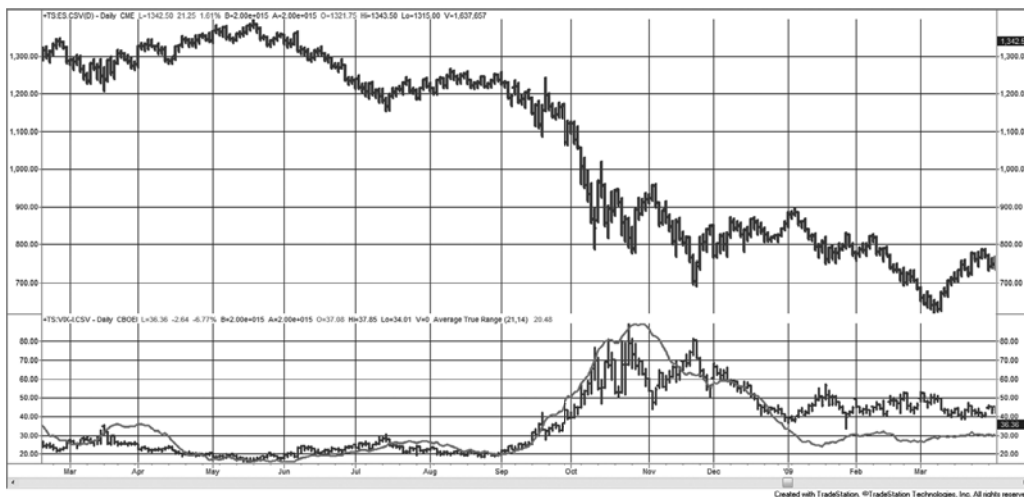
FIGURE 20.7   *e*Mini S&P (top) and the CBOE VIX index with the average true range in the bottom panel.

was trading near its lows of about 20. The VIX pattern, which is generally inverse to the S&P, leads to the belief that the S&P rises on low volatility and falls on high volatility. Another way of seeing this is that investors are exiting on high volatility and entering on low volatility. For that reason traders talk about buying the S&P when VIX is high and selling when it is low.

For comparison, Figure 20.7 also shows the historic volatility in the lower panel. This is actually the 21-day average true range of the *e*mini S&P, but it tracks the implied volatility closely except when prices are under stress. Remember that the implied volatility is what options traders believe will be the volatility *over the next 30 days*, while historic volatility is the actual market volatility *over the past 30 days* (actually 21 trading days). Note the key difference is that, right after the October decline the implied volatility dropped faster than the actual market volatility, then the implied volatility spiked again in late November when the actual volatility continued to fall.

USING VOLATILITY FOR TRADING

For many traders, volatility is essential to success. It has been embedded in many of the systems and methods discussed throughout this book, but this section will look at specific applications that use volatility.

Intraday Volatility and Volume

In Chapter 12, there is an intraday pattern showing that volume is highest near the opening of the day and again near the close of trading. Intraday volume forms a smooth *U*-shaped

pattern for domestic markets, and a similar but more complex pattern for markets that open in Europe but span U.S. trading hours.

Intraday volatility has a pattern that is identical to volume, highest at the open, then declining to its lowest point at mid-session, and rising again as the trading day ends. The closing volatility and volume are typically lower than the opening activity.[3] For this purpose, volatility (*VT*) was calculated

$$VT_t = \sqrt{\frac{1}{n-1} \sum_{i=t-n+1}^{t} \left(\ln\left(\frac{p_i}{p_{i-1}}\right) - \bar{P} \right)^2}$$

where n = calculation period
  ln = natural log
  p_i = price at time i
  $\bar{P}$ = arithmetic mean of the natural log of price differences:

$$\bar{P} = \frac{1}{n} \sum_{i=t-n+1}^{t} \ln\left(\frac{p_i}{p_{i-1}}\right)$$

To find the correlation between the intraday pattern of volatility and that of volume, a simple linear regression can be solved

$$Volatility_t = a_t + b_t \times Volume_t + e_t$$

where all values are calculated at time t. The resulting correlation, $R = 0.595$, is statistically significant for NASDAQ 100 volatility and volume. The correlation between volatility and volume are highest at the beginning and end of the day. Meissner and Cercioglu suggest that this volatility pattern, with the corresponding volume that provides liquidity, can be traded by being long options at the beginning and end of the day, profiting from the *gamma* (the rate of change of *delta*, which is the rate of change of the futures price with respect to the rate of change of the underlying asset). During the quiet mid-session period, a short options position may be used to profit from *theta*, the time decay.

Predicting Volatility with Trading Ranges

William Brower, in his publication *The Inside Edge*,[4] performed a thorough study of trading ranges, looking for when there would be a higher volatility tomorrow. This is an advantage for day traders and systems with short holding periods, and breakout systems in general. A summary of his results is shown in Table 20.1 for the S&P from 12/23/87 through 12/15/95. *O*, *H*, *L*, and *C* are the opening, high, low, and closing prices, and the notation [1] and [2] represent the prices 1 and 2 bars back.

---

[3] Gunter Meissner and Sarp Cercioglu, "Trading Trends and Correlations," *Futures* (January 2003).
[4] William Brower, *Inside Edge* (Inside Edge Systems, 24 Old Stagecoach Road, Redding, CT 06896, March/April 1996).

TABLE 20.1 Predicting the Trading Range of the S&P 500

Rule	Conditions Tested	Conclusion
1	`0 < L[1] − x`	Very good predictor, but few cases (161 at $x = -0.10$)
2	`0 > H[1] + x`	Modest predictor, but few cases (101 at $x = 0.60$)
3	`C[1] + x > 0 > C[1]`	Range got smaller when $x = 0.35$
4	`C[1] − x < 0 < C[1]`	Range got smaller when $x = 0.30$
5	`0 < C[1] − x`	Range tended to increase as x increased
6	`H[2] <= H[1] and L[2] >= L[1]`	No significance
7	`H[2] > H[1] and L[2] < L[1]`	Modest predictor of lower volatility
8	`Day of week`	Mon had lowest volatility, Tues and Fri the highest
9	`Average(TrueRange,3)[1] > x`	Higher average true range for 3 days was a very good predictor of higher volatility
10	`RSI(close,3)[1] < x`	When RSI < 40 good predictor of higher volatility

Source: William Brower, *The Inside Edge.*

Thomas Bierovic—On-Balance True Range

To visualize the change in volatility, Thomas Bierovic has created an *On-Balance True Range* by following the same rules as On-Balance Volume (*OBV*), but substituting the true range calculation for volume. He then calculates a 9-day exponential smoothing of the On-Balance True Range and uses the crossovers of the oscillator and smoothed oscillator to confirm signals. Although the highs and lows may come at nearly the same time as other oscillators, the relative peaks and valleys may offer the trader new insights. For many traders, this simple interpretation can help separate high and low volatility conditions.

VIX Trading Systems

The *VIX*, the index of implied volatility (see details in the previous section), can also be interpreted as the level of risk in the market. VIX is considered a mean-reverting indicator as the following strategies show.

Connors

Larry Connors has based a number of trading systems on the VIX. This strategy, called *Connors VIX Reversal 9, CVR* 9, is based on VIX range expansion, looking for higher prices in the S&P Index following a VIX expansion and lower prices following a VIX contraction.[5] Connors treats volatility as mean-reverting. Entries are based on a minor reversal in the VIX. The rules for buying (selling are the reverse) are:

1. Today's VIX high must be higher than the VIX high of the past 10 days.
2. Today's VIX must close below its open.
3. Yesterday's VIX must have closed above its open.

---

[5]Larry Connors, "Timing Your S&P Trades with the VIX," *Futures* (June 2002).

4. Today's VIX range must be greater than the ranges of the past 3 days.

5. If conditions 1–4 are met then *buy* S&P futures on the close and *exit* in 3 days.

Connors is actually looking for turning points in the VIX. The specific pattern that precedes a buy signal in the S&P ends with a range expansion. This expansion is likely to mark the end of a short-term upwards move in the VIX. A decline in the VIX that follows eases the way for a short-term rally in the S&P. Traders are more comfortable buying when volatility is dropping, or at least not at noticeably high levels. As with many other systems, this requires additional protective stops and position size management. A program for testing this, *TSM VIX Connors CVR9*, is available on the Companion Website.

Another Connors mean-reverting VIX strategy[6] uses the RSI for timing. If

1. The S&P > 200-day moving average

2. The 2-day RSI of the VIX > 90

3. Today's VIX open > yesterday's VIX close

then *buy* on the close and *exit* when the 2-day RSI closes > 65.

MarketSci Blog

One of the more interesting websites is *MarketSci Blog*, which offers numerous creative strategies for equities trading. This one, published on March 1, 2011, takes a 10-day exponential smoothing (*EMA*) and a 10-day simple moving average (*SMA*), both applied to the VIX index, and *buys* the VIX when the EMA falls below the SMA, then *sells short* when the EMA moves above the SMA. It does this based on the concurrent closing prices. The rationale is that the EMA is faster than the SMA, given the same calculation period; therefore, it acts as a timing trigger. A fast execution is essential for mean-reverting trades. Figure 20.8 shows the VIX prices and the two trends for April to October 2010.

Figure 20.9 breaks down the profits from the VIX mean-reverting strategy into long only, short only, and both long and short. Unlike most other strategies, this is remarkably symmetric, with longs and shorts performing equally. It is important to note that this study was done using the VIX cash index, which is not tradable. There is a futures contract for VIX, but not enough data for a valid study.

Gerald Appel on VIX

Gerald Appel[7] gives his own thoughts on trading VIX as

- *Buy* when there are high levels of VIX, implying broad pessimism.
- There are no reliable sell signals using VIX.
- Volatility tends to increase during weaker market climates.
- The stock market is likely to advance for as long as volatility remains stable or decreasing.

---

[6] Connors, Larry, *Short-Term Strategies That Work*. (Jersey City, NJ: Connors, 2008).

[7] Gerald Appel, *Power Tools for Active Traders* (Upper Saddle River, NJ: Financial Times, Prentice-Hall, 1999), 155.
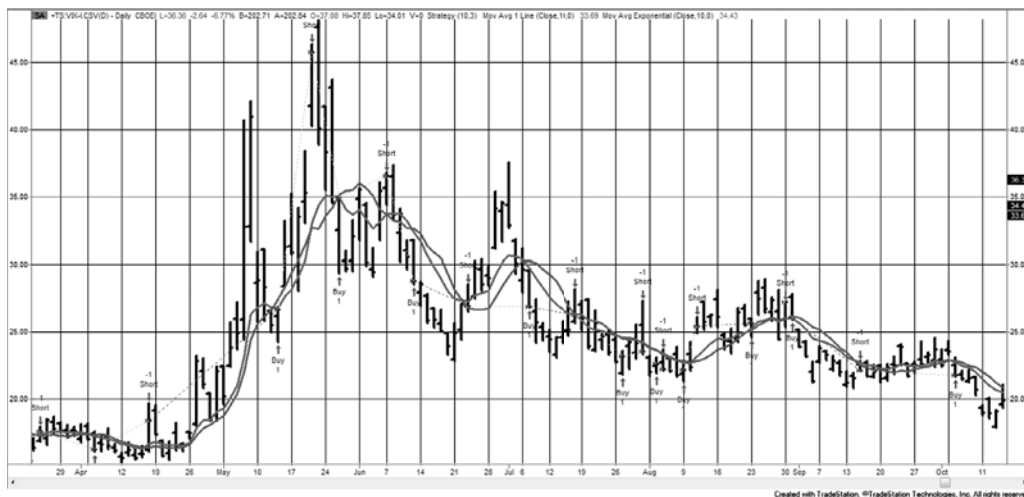
FIGURE 20.8 VIX mean-reverting strategy from MarketSci Blog, triggered by the EMA crossing the SMA.

Volatility System

Bookstaber[8] uses the average true range (*ATR*) over the past n days as the basis for a simple volatility strategy:

- *Buy* if the next close, C_{t+1}, rises by more than $k \times ATR_t(n)$ from the current close C_t
- *Sell* if the next close, C_{t+1}, falls by more than $k \times ATR_t(n)$ from the current close C_t
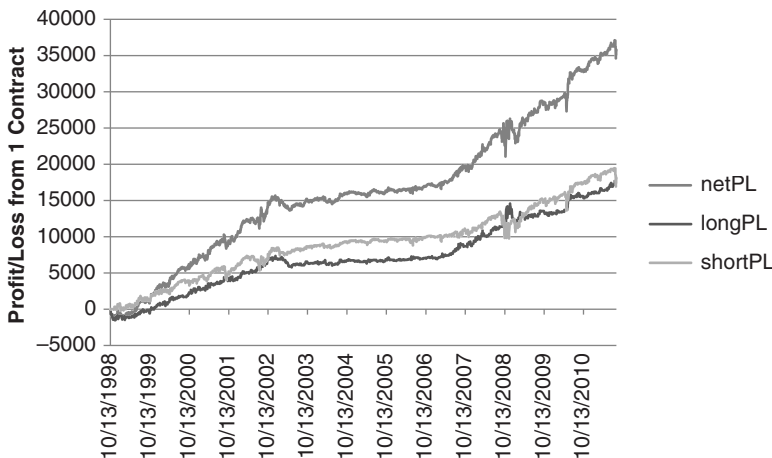
FIGURE 20.9 Profits from VIX Mean-Reverting Strategy, 1998–2011.

---

[8]Richard Bookstaber, *The Complete Investment Book* (Glenview, IL: Scott, Foresman, 1985), 224–236.

The volatility factor k is given as approximately 3, but can be varied higher or lower to make the trading signals less or more frequent, respectively. This method falls into the category of *volatility breakout*.

Profit Targets and Stop-Loss Orders

One of the most common uses of volatility is to make both profit-targets and stop-loss orders adapt to current market conditions. For that purpose, a method identical to Bookstaber, in the previous section, is most often used to define both; however, they are often applied differently.

Profit targets are best used for short-term trading and determined at the time of entry

$$PT_t = p_E \pm k \times ATR_{t-1}$$

where PT is the profit target, p_E is the trade entry price, and k is a constant, usually about 3. If the last entry was long, the target is added, and when it is short, the target value is subtracted from the entry price. The profit target is triggered by the intraday high or low price, rather than the close. The reason that profit targets are used most often in short-term trading, rather than in long-term trend following, is that shorter time intervals have a higher degree of market noise. Once profits are taken, you want prices to reverse and allow you to reenter at a better price, or reverse direction entirely. For long-term trend following, it is necessary to capture the *fat tail*; therefore, taking profits works counter to that notion. Once you exit a long-term trend by taking profits, and the trade continues to be profitable, it is necessary to reenter the trade to avoid losing that rare chance at an exceptionally large profit.

For *stop-losses*, the same $k \times ATR_{t-1}$ is used, but it is subtracted from the long entry and added to the short entry. In this case, when applying stops in a noisy market it is best to use the high and low to trigger the stop, but actually exit the position on the close. That way you take advantage of the market noise to improve the exit price. Stop-losses have the same problem as profit-targets when applied to long-term trend systems. If you are stopped out, but the trend is still intact, it is necessary to reenter the trade in order to avoid missing the rare but large profit.

TRADE SELECTION USING VOLATILITY

High volatility is clearly related to greater risk, but low volatility may also mean that there is a smaller chance for profits. The following are reasonable expectations for selecting trades based on volatility:

- *Entering on very high volatility* is exposure to very high risk. Returns from high volatility trades may range from large profits to large losses. The net results may be a profit, but the return-to-risk ratio is likely to decline. Long-term performance may be best if these trades are avoided.

- *Entering on extreme low volatility* seems safe, but prices often have no direction and produce small, frequent losses. Waiting for an increase in activity before entering might improve returns. Contrary to this, some traders advocate entries when prices have had a short-term drop in volatility.
- *Exiting a position when prices become very volatile* should reduce both profits and risk, but may come too late. This is an issue best resolved by evaluating the specific system.

Eliminate or Delay?

Whenever a high or low volatility situation occurs at the time of an entry signal, there are two choices. The trade can be completely eliminated by *filtering*, or it can be delayed until the high volatility drops or the low volatility increases to an acceptable level. When a trade is filtered, it is necessary to track the trade that was not entered, to know when it was completed. Each new trade is subject to the volatility threshold at the time of its entry. A trade that is delayed pending a change in volatility can be entered any time volatility moves back into the acceptable range. The following sections look at these cases in more detail. Before starting, we can theorize that short-term trading would most likely eliminate, not delay, trades that fall outside the acceptable volatility range because there are many trades and each is held for a short time. At the other end of the spectrum are the long-term trend trades, held for weeks, that would suffer if the exceptionally large profit was missed. Then delaying the entry would be the best solution.

Constructing a Volatility Filter

Calculating the volatility is simple to program using any spreadsheet or strategy testing software. The following steps were used here:

1. Calculate a moving average trend. Use one fast trend and one slow trend.
2. Calculate the volatility, using any one of the methods described early in this chapter, but not including the volatility of the current day.
3. Enter a new trade (on the close) if today's volatility is (a) above the low filter threshold or (b) below the high filter threshold.
4. Exit a current position if the volatility is above the high filter threshold *and*, based on testing, (a) the current price change has moved in a profitable direction or (b) the current price change has moved in a losing direction.

These six different conditions will be applied to five very different markets: Eurodollars, Japanese yen, crude oil, IBM, and the S&P 500 over the period 1990 through mid-2011 to allow enough trades for the long-term trend to show valid results. We know that this interval contained exceptional and varied price moves.

The two trends used will be 10 and 60 days. A charge of $15 per trade per side will be used so that there will be some impact of trading more often. It is likely that some

of these markets will not be profitable using these trends, but the results are more concerned with improvement rather than absolute profits.

Standard Deviation Measurement

A standard deviation was used to determine the volatility threshold level because these levels are associated with probabilities. A high-volatility filter with a *1-standard deviation threshold* means that no trades were taken if the volatility was above the average volatility plus 1 standard deviation, the top 16% occurrences. A *2-standard deviation threshold* filters out volatility in the top 2.5%, and 3 *standard deviations* restricts only the top 0.13%. In all cases, a 20-day standard deviation will be used, comparable to VIX. Because the period is short, volatility can jump well beyond the normal 3-standard deviation maximum, and it may be necessary to test filter values above 3.

The program *TSM Moving Average* will used as the underlying strategy. This enters and exits trades based entirely on the direction of the moving average trendline, not on the price penetration of the trendline.

Entry Filter Results

Table 20.2 shows the results of a 40-day moving average system (a moderately slow trend), buying when the trendline turns up and selling when it turns down, using a volatility entry filter that delays the entry until the threshold conditions are met. The data used was back-adjusted futures from 1990 through July 2011, except for Amazon, which began in May 1997. Costs were $15. Results are expressed as total profits or losses, profit factor (gross profits divided by gross losses), number of trades, percentage of profitable trades, and profits per contract or per share (for Amazon). Each market has three lines, the

TABLE 20.2 Tests of Low and High Threshold Filters, from 1990 through July 2011, Except for Amazon, Which Began in May 1997

Market	Test	$PL	Profit Factor	Number of Trades	% Prof Trades	Profits/ Trade
Eurodollars	No filters	41,087	2.11	247	31.5	166
	< 0.25	42,537	2.5	175	34.5	244
	> 1.60	42,425	2.56	158	37.4	268
Euro FX	No filters	100,062	1.33	302	35.4	331
	< 0.30	91,225	1.36	233	37.7	391
	> 0.55	102,975	1.36	282	36.5	365
Crude oil	No filters	121,450	1.49	323	36.8	376
	< 0.35	135,170	1.67	246	41.9	549
	> 2.55	126,560	2.11	105	41.9	1205
S&P	No filters	−37,187	0.83	378	28.3	−98
	< 0.05	6,650	1.07	153	29.4	43
	> 3.0*	3,825	1.06	88	30.7	43
AMZN	No filters	58.41	1.16	190	32.6	0.31
	< 0.15	124.10	1.51	115	36.5	1.08
	> 1.30	96.09	1.33	133	37.6	0.72

first showing the results of the moving average system with no entry filters, the second reflecting only trades below the low filter, and the third only trades above the high filter.

The low filters shown on each of the second lines in column 2 were found by optimizing the filter times the standard deviation of returns over the past 20 days over the range from 0.05 to 1.0 in steps of 0.05. Eurodollar interest rates show that the optimal threshold was 0.25, which reduced trades by 30% and increased profits per trade by 47%. All of the markets show significant increases in performance when only low volatility trades were taken. In the case of the S&P, a losing moving average performance will normally be improved by removing trades. All markets show improvements when the threshold level is below 0.35.

With the high volatility filter, which allows trades only above its threshold level, we expected to show profits but at higher risk; however, it did better than expectations. Test inputs were 0.25 to 3.0 in steps of 0.05. The S&P was the only market that chose the maximum filter value, trying to remove as many trades as possible. Other markets all showed good improvement, but the threshold levels were not consistent, depending on the nature of price movement. Crude oil, which has had the most volatility and the largest price moves, shows best profits with a threshold at 2.55, a level that would eliminate most trades in other markets.

Figure 20.10 shows the net equity from the Eurodollar interest rate futures using the high and low filters separately. Results are similar because the 20-day standard deviation is highly variable and both filters will eventually allow an entry to the same trade, but at different places. As we can see from the chart, the low volatility entries outperform the high volatility and produced a smoother return stream. If we calculate the information ratios for the two tests, rather than use the profit factor, we see that the low-volatility filter has a return of 6.08% with a volatility of 5.32% while the high filter has a return of 4.85% with volatility of 5.59%. That gives the low filter a ratio of 1.143 compared to 0.867 for the
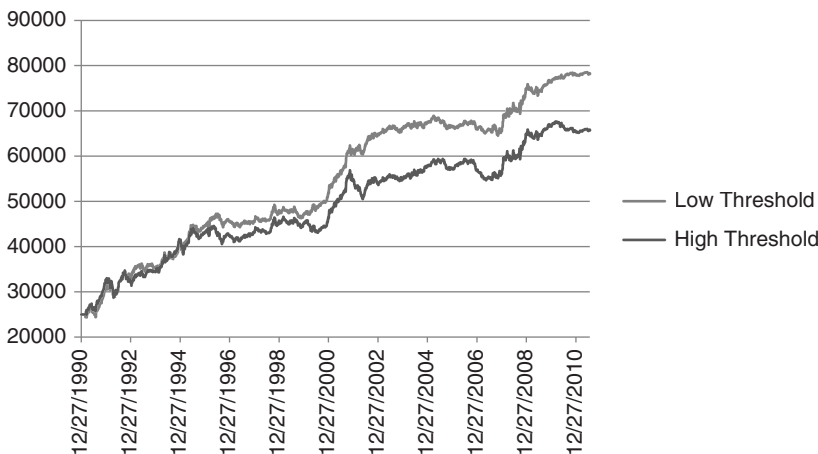
FIGURE 20.10 Net equity resulting from Eurodollar rates using a 40-day moving average and a low filter of 0.25 compared to a high filter of 1.60.

high filter, considerably better. These calculations can be found in the Excel spreadsheet, *TSM Moving Average Filtered*, along with the strategy, on the Companion Website.

This test confirms the popular notion that entering on low volatility is preferable. Even without any high-volatility trades, this seems to be a safe approach. The program *TSM Moving Average Filtered*, available on the Companion Website, was used for testing.

High Volatility Exits: Reducing the Risk

It seems natural to think that you may be able to reduce risk by exiting a trade when volatility becomes very high. It also may be too late because you may have already been exposed to the worst of the volatility. If prices jumped because of a price shock, then the highest volatility would have occurred at the moment of the shock. On the other hand, if prices have been moving higher for weeks, and are becoming very volatile at the top of a move, then exiting could save a lot of aggravation. If a market is noisy, then volatile price jumps are frequent and short-lived. In those situations, you should close out positions when a volatile move is profitable because profits will soon disappear. Or, if the volatile move is an immediate loss, then waiting before you exit should recover at least part of the loss. The opposite is true for a trending market. A good price move should be followed by more profits, and a bad move by further losses.

A high-volatility exit is not the same as profit-taking because it is not targeting a specific profit for the trade, but looking only at the risk of a 1-day move. One of the more difficult tests should be U.S. 30-year bonds with a long calculation period because profits are normally the result of holding onto the trade until the end, capturing the fat tail. Nevertheless, we will test high-volatility exits with the following rules:

1. Use a 40-day moving average and generate buy and sell signals when the trendline turns up and down.

2. Calculate the 20-day volatility as the standard deviation of the 20-day price returns.

3. Exit the current trade, whether long or short, when today's returns are greater than the previous 20-day volatility times a factor.

4. If a high-volatility exit occurs, do not reenter until the trend changes.

Based on bond futures prices from 2000 through July 2011, Table 20.3 shows the results. Line 1 is the benchmark performance of the 40-day moving average without high-volatility exits. Line 2 has the same entries but exits when the returns are 10 times the

TABLE 20.3 Test Results for High-Volatility Exits, 30-Year Bonds, 2000 through July 2011

Test	Net PL	Profit Factor	Number of Trades	% Prof Trades	Profit / Trade
No HV exits	4,500	1.02	350	36.0	12
Exit factor 10	12,968	1.07	350	40.0	37
Exit and reenter	34,781	1.16	387	40.8	89

20-day volatility. Results are somewhat better, but then only the most extreme moves were used to exit.

Remembering that trend systems need to capture as much of the very good moves as possible, we will add a *reset rule*. We will be able to reenter the market in the same direction as the trade just exited, but only when the volatility drops to one-quarter the volatility (the daily return) on the day of the exit. The concept is that, at one-quarter the volatility, the market is more-or-less back to normal. Results of that test are shown on line 3 and are considerably better. A program to test this is *TSM Moving Average Volatility Exits*, available on the Companion Website.

Ranking Based on Volatility

Gerald Appel[9] offers an additional approach to trade selection by creating a ranking method for mutual funds.

- Select only funds with average to below average volatility.
- Add the 3-month and 12-month performance together to get a single value.
- Rank the funds.
- Only invest in the top 10%.

Trade Selection Summary

Once you have a basic system, the next step is to decide whether you can eliminate some of the losing trades without eliminating the profitable ones. In the extreme, we want to eliminate *all* of the losing trades, even if we reduce some of the profits; of course, that is impossible. Every system has a risk, even so-called riskless trades. Arbitrage, when done properly, has virtually no risk; however, it may be so competitive that the opportunities are rare and the margin of profit small. In the final analysis, you can't remove the risk, only delay it or move it around. If your system shows that it has essentially no risk, then it's important to rethink your development process to find the flaw.

Nothing is free. Good trading is not a matter of hitting it rich on a single short position, it is grinding out a profit day by day and week by week. Occasionally, if you follow new markets, you might find an opportunity that others have not yet seen. If you act fast, you can capitalize on this for a short time, until others see the same situation. Then bigger players come in and push you out. In turn, they are followed by other investors who combine to remove all the profit opportunity, even for themselves. At the same time, other opportunities surface.

When selecting trades to eliminate, the easiest place to begin is by associating performance with volatility or price level. While some systems perform better in an environment of higher volatility, your strategy may show the best return relative to risk when there is less volatility. When you filter trades you will always get rid of good ones while, hopefully, removing more of the bad ones.

---

[9] Gerald Appel, *Power Tools for Active Traders* (New Jersey: Financial Times, Prentice-Hall, 1999)

LIQUIDITY

Many systems that have excellent historic backtests fail when they are actually traded. These include models based on both intraday and daily data. One reason for this disappointing performance is the lack of understanding of market liquidity at the time of execution. Consider two systems:

1. A *trend-following method*, which will trigger buy or sell orders as prices rise or fall during the day.
2. A *countertrend system*, which sells and buys at relative intraday highs or lows.

Although each system intends to profit with their opposing philosophies, both act when prices make a relatively unusual move. Figure 20.11 shows the normal distribution of intraday volume and the profitability associated with this distribution.[10]

The solid line *bb'* is the actual volume on an ideally normal day. Volume is greatest at the median *M* and declines sharply to the high and low endpoints *H* and *L*, where only one trade may have occurred. As we know from previous sections, the time of highest volume will be near the open and close of trading; therefore, *M* will actually occur at the beginning or end of the day, while the end points of *bb'* will occur near midday.

The dotted line *cc'* represents the apparent profit for a countertrend system that makes the assumption of a straight-line volume distribution. The endpoints are shown to contribute the largest part of the profits when, in reality, no executions may have been possible near those levels. Assuming the ability to execute at all points on the actual distribution *bb'*, the approximate profit contribution is shown as *dd'*.
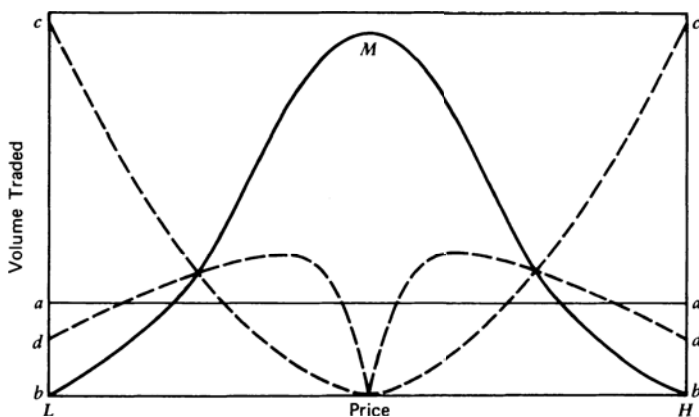
FIGURE 20.11   Daily price/volume distribution shows liquidity.

---

[10] Gary Ginter and Joseph J. Ritchie, "Data Errors and Profit Distortions," in Perry J. Kaufman, ed., *Technical Analysis in Commodities* (New York: John Wiley & Sons, 1980).

For trend-following systems, no profits should be expected when buy or sell orders are placed at the extremes of the day. The actual price distribution bb' is the maximum that could be expected from such a system; in reality, the first day is usually a loss.

TRENDS AND PRICE NOISE

Throughout this book there are techniques that try to find the underlying trend when the price direction is not clear. The trader who enters a new trend sooner will be more profitable, but the erratic behavior of prices makes a faster response less reliable. What makes a trend so difficult to identify is *noise*. Noise is the erratic movement of price and, by definition, it is unpredictable. In engineering, this type of behavior that shows no patterns is called *white noise*, yet the noise that we see in price movement has very definite statistical attributes. In Chapter 1 we discussed the way to measure noise using the efficiency ratio, or fractal efficiency. We found that different markets had different levels of noise, but that the equity index markets had the most noise and short-term rates the least. Also, developing markets had low noise while markets in the major industrialized countries had more noise. As the financial systems mature, more investors participate, and noise increases. In this section we will look at another aspect of noise, how it directly affects trading. It is still treated as a big picture item, more important conceptually.

Noise is the product of a large number of market participants buying and selling at different times for different purposes. Each has its own objectives and time frame. Noise can be the result of price shocks—unexpected events, in particular news or economic reports, that causes change that persists for varying periods into the future. Noise has most of the qualities of a sequence of random numbers. Nearly half of all price moves change direction on the next day; about 25% of all price moves continue in the same direction for two days, 12.5% for three days, and so on. These patterns should already be familiar to analysts. The size of a price move, or price shock, also appears to be random; that is, 50% are quite small, 25% are twice as large, 12.5% are four times as large, and a few are extremely large.

Then how can you tell if a price move up is an indication of a new direction or simply more noise? The answer is that you cannot know for sure until prices continue in the same direction. If you are a fundamental analyst, a structural change in the market is as close as you get to a definite change. For example, the Fed continued to lower interest rates following the stock market peak in 2000 in order to stimulate economic growth and prevent a deep recession. By the end of 2003, equity prices had been rising for six months and interest rates were at a 50-year low. A move by the Fed that raises rates would be a structural change, a clear sign that it was more concerned with the prospects of inflation and convinced that that economic growth would continue without further stimulation. A trader would expect that a short position in long-term maturity, fixed income futures, would be sound.

A systematic trader, studying only the numbers, could identify this specific structural change by following the government's lending rate, the prime rate, or a combination of short-term rates, in combination with the price of market being traded. But at this point we are only concerned with how to recognize a trend change looking at a single price series.

Figures 20.12 a–c show how noise varies for different markets and how it affects trend trading. The first is Eurodollar interest rates, typical of a very low-noise market. Because of less noise, a simple breakout strategy would be able to identify a change of direction in time to capture a large percentage of the price move. The second part represents a market with more noise, but still identifiable trends, such as the EURUSD. A higher level of noise means that the trend is identified later and the exit later, resulting in smaller profits for about the same size move as Eurodollars. The third figure represents the equity index markets, which have exceptionally large noise. By the time the new trend can be identified, entered, and finally exited, there is no profit.

We can improve our chances of entering a new trend, and not being fooled by noise, by waiting until the price changes direction by a larger amount. A small price change could
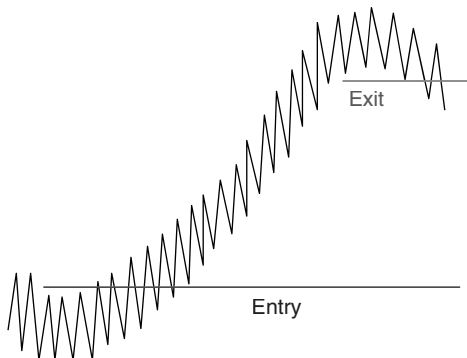
FIGURE 20.12a   A low noise market allows sooner entries and exits resulting in larger per trade profits.
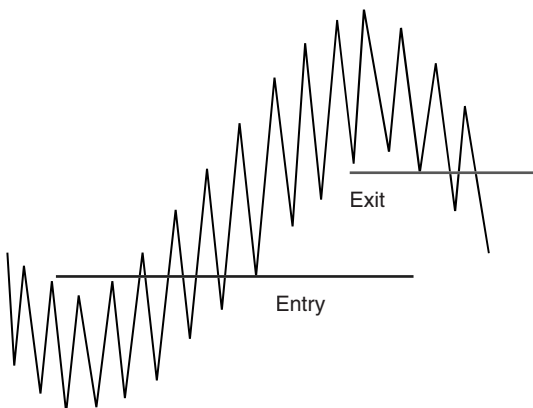
FIGURE 20.12b   Increased noise causes entries and exits to be delayed, resulting in less profit for the same-size move.
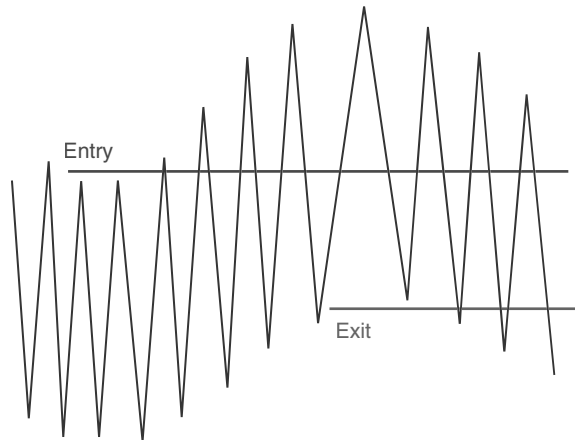
FIGURE 20.12c   High-noise markets, typical of an equity index in a major industrialized country, cause entries and exits that are too late to net a profit.

easily be noise, but as the price change gets larger, the chance of being a true trend increases quickly. This is the reason why a long-term trend is more reliable than a short-term one.

Figure 20.13 tries to visualize the way noise and the trend calculation period work together. Along the bottom are the absolute returns, showing frequent "normal" returns and occasional price shocks, some smaller, some larger. When a fast trend period is used, it experiences a large number of price shocks that might cause the trend
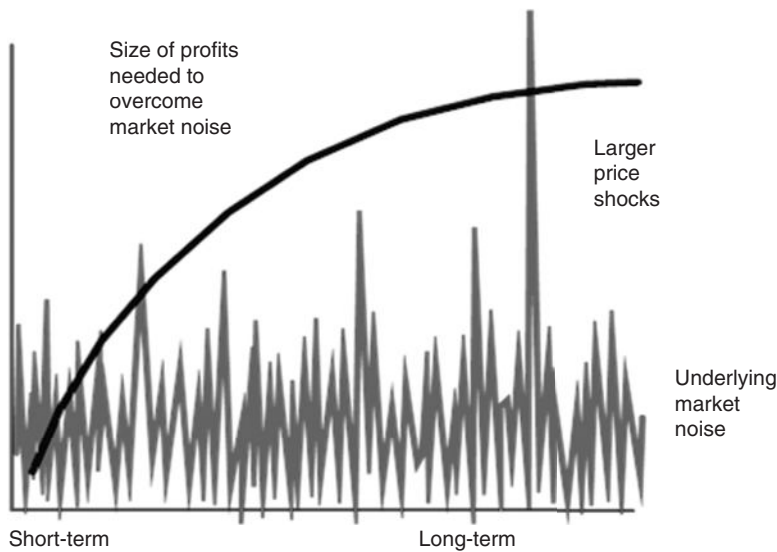
FIGURE 20.13   How the trend calculation period (the curved line) relates to market noise.

to change. As the trend period gets longer, the trend lags behind, and fewer shocks will trigger a trend change. This results in the longer trends being more reliable, although they will suffer from markets swings while holding onto the same trade. There is always the exceptional price shock that affects any trend, as shown toward the right part of the illustration.

As markets evolve, noise increases. To capture the same returns, it may be necessary to use a longer calculation period for trends. If you are already doing that, then expect smaller returns. One alternative is to take advantage of the noise and use mean reversion methods, emphasizing the shorter holding periods. Noise is dominant in the short term and trends in the long term.

TRENDS AND INTEREST RATE CARRY

In the environment of exceptionally low interest rates during 2010–2012, an interesting new situation has developed expressed by the question "Can you profit from using a trend system for interest rates when rates are near zero?" Certainly, there seems to be a good reason for believing that there is little upside profit trading 10-year notes and much greater downside risk if yields suddenly increase. Yet, this is not really the case.

Trading interest rate futures involves rolling into the new contract every three months ahead of expiration. Each time the data rolls into the next delivery month, it shifts from about 15 days to expiration to at least 105 days to expiration. The net interest on 105 days is much greater than 15 days, so the price of the contract will be lower because it is either discounted (as in the case of short-term rates) or simply reflecting a higher yield-to-maturity. If this situation persists for many months, or even years, then staying long the interest rates continues to extract the carry. While that may not be a large profit in any one contract, the accumulated returns can be significant.

EXPERT SYSTEMS

An *expert system* is one in which you draw conclusions based on an accumulated *knowledge base* that has been stored as data, facts, and relationships in the form of *if-then-else* rules. The person who collects this information is called a *knowledge engineer*. This process is often shown as a decision tree, where each "yes" or "no" answer leads down the binary path to the final conclusion; however, the actual process is more sophisticated and more interesting. Unfortunately, the past years have not seen new information published on trading systems using this technology. It is generally thought that any successful solution would be held confidential.

For a trading strategy, data used in the knowledge base should include price and economic data as well as the knowledge that, for example, *highly correlated markets*

move together, and *high volatility means high risk*. When the first expert systems were developed, the core information was actually gathered by interviewing experts, which is how the name *expert system* was derived. The success of this method depends upon the quality and completeness of the knowledge base.

The knowledge base is used by an *inference engine*, which is able to draw conclusions from the *facts* stored in the knowledge base. Therefore, if we have the following information,

FACT   1: U.S. bonds are more volatile than Eurodollar rates.
FACT   2: The S&P 500 is more volatile than U.S. bonds.
FACT   3: Volatility is directly proportional to risk.

then the inference engine can create the new fact

FACT   4: The S&P is riskier than Eurodollars.

The inference engine provides a straightforward, logical process; however, there may be many relationships to resolve, not all of which may apply to the problem you would like to solve. For example, if we also have the following facts:

FACT   5: The S&P has a high degree of noise.
FACT   6: Eurodollars have a low degree of noise.
FACT   7: The S&P is currently trading below its 200-day moving average.
FACT   8: The S&P is currently trading above its 20-day moving average.
FACT   9: The S&P has been above its 200-day moving average for 80% of the past 20 years.
FACT 10: Eurodollar rates are driven by monetary bank policy.
FACT 11: Current monetary policy is dominated by concerns of inflation.
FACT 12: Inflation results in higher interest rates.
FACT 13: Net trading returns for the S&P have been 40% per annum.
FACT 14: Net trading returns for the Eurodollar have been 20% per annum.

then human logic might conclude that Eurodollar yields are likely to rise because of concerns over inflation. In addition, the S&P has greater risk, more erratic behavior, and is currently not as strong as it has been on average over the past 20 years, although it is now rising. Compared to Eurodollars, there is greater risk trading the S&P, but its returns have been higher.

How can this expert conclusion be duplicated by a computer? By adding a set of rules that parallels the thinking of experts, a computer can theoretically arrive at the same conclusions. For example,

RULE 1a: IF the Producer Price Index rises by more than the annualized rate of 4%, THEN we have inflation.
RULE 2a: IF there is high noise, THEN there is less chance of a trend.
RULE 3a: IF there is high noise, THEN there is greater risk.
RULE 4a: IF the faster trend is above the slower trend, THEN prices are trending up.

For each positive rule 1 through 4, there should also be a negative rule:

RULE 1b: IF the Producer Price Index does not rise by more than 4% annualized, THEN we do not have inflation.
RULE 2b: IF there is low noise, THEN there is a greater chance of a trend.
RULE 3b: IF there is low noise, THEN there is lower risk.
RULE 4b: IF the faster trend is below the slower trend, THEN prices are trending down.

Even with the negative rules there are some ambiguous cases. For example, in Rule 4 two trends are in conflict. In other situations, the positive rule might be true, but the negative may not be as strong. In Rule 1b, we see that the effect is that "we do not have inflation" rather than "we have deflation."

Forward Chaining

The process of combining the rules and facts in order to yield an expert opinion is called *forward chaining.* For example, beginning with FACT 6, "Eurodollars have a low degree of noise" we find the relevant rule, RULE 3b: "IF there is low noise, THEN there is lower risk" and create a new fact that "Eurodollars have (relatively) low risk." Note that in each case the terms *low*, *high*, and *faster* are all relative.

Once we have this new fact, that Eurodollars have relatively low risk, and we similarly conclude that the S&P has "relatively high risk," we can also conclude that the Eurodollars have lower risk than the S&P. This process of following the path of each fact as it is handled by various rules is the forward chaining. It will lead to other rules and other facts; it may be that the expert opinion will be found along this route, or that the combination of new facts, such as the relative risk between the Eurodollars and the S&P, will provide the expert opinion.

This example shows only a few facts that can be easily summarized; however, there are thousands of pieces of information about performance characteristics, relationships to other markets, and fundamental factors that might alter expectations. If accumulated by asking experts, it is also likely that there will be conflicting information. While the human brain has a remarkable ability to sort through these items and select the information it considers most relevant, some important items can be overlooked when there is too much to consider. An expert system is expected to use all of the data and reduce it to a single decision. In doing this, it must also select the most significant facts and resolve conflicts associated with the proper order of events and the time horizon of the investor.

A Technical Expert System

An expert system can treat indicator values as expert opinions and create systems with the fundamental relationships shown in the previous section. An example by Fishman,

Barr, and Loick[11] applied to the DJIA defines the *rules* as the relationships between the various indicators and calculations, and the *facts* as the values of those items. The purpose of this expert system is to inspect trending and nontrending characteristics of price movement to give the probability of a continued trend. The following example is adapted from their article:

Rule 1: IF ADX > 18 AND ADX >= ADX[2]
        AND nontrending is false or undefined
        THEN there is a 95% chance the market is trending
Rule 2: IF ADX < 30 AND ADX <= ADX[2]
        AND trending is false or undefined
        THEN there is a 90% chance the market is nontrending
Rule 3: IF the market is trending
        THEN MACD = probability of 0.8 AND SD = probability of 0.85
Rule 4: IF the market is nontrending
        THEN MACD = probability of 0.55 AND SD = probability of 0.75 (75%)
Rule 5: IF stochastic AND SK[1] < 30 AND SD[1] < 30 AND SK[1] < SD[1]
        AND SK < SD THEN sell with 80% confidence
Rule 6: IF stochastic AND SK[1] > 70 AND SD[1] > 70 AND SK[1] > SD[1]
        AND SK > SD THEN buy with 80% confidence
Rule 7: IF MACD AND MACD value > signal line value
        THEN buy with 75% confidence
Rule 8: IF MACD AND MACD value < signal line value
        THEN sell with 75% confidence

where   ADX = the current 18-period ADX calculation
        ADX[2] = the ADX value 2 periods ago
         MACD = the Moving Average Convergence/Divergence with smoothing con-
                 stants 0.15 and 0.075
           SK = a 9-period, %K fast stochastic
           SD = a 9-period, %D slow stochastic

Undefined means that there is no information about the values. The expression "*IF stochastic*" is a test to see if this option is being used. The *facts* are the actual values to be used in the sample evaluation:

FACT 1: ADX = 19
FACT 2: ADX[2] = 19

---

[11] Mark B. Fishman, Dean S. Barr, and Walter J. Loick, "Artificial Intelligence and Market Analysis," *Technical Analysis of Stocks & Commodities* (Bonus Issue, 1993).

FACT 3: SK = 90
FACT 4: SD = 80
FACT 5: SK[1] − SK = 68
FACT 6: SD[1] − SD = 92
FACT 7: signal line value = −70
FACT 8: MACD value = −68
FACT 9: the market is trending

The results, expressed as probabilities, offer greater insight into the likelihood of success using this method. Arriving at these probabilities, however, requires additional decisions. Without any further information, we can assume that there is only a 50% chance that the stochastic is a correct trend indicator. If we were to test the number of times the stochastic indicated an upwards move with the number of days that the subsequent price moved higher, we could get a much better indication of the chance of success.

FUZZY LOGIC

In the second example of expert systems, the results were expressed as probabilities. Because trend and indicator calculations are estimations of market movement, very little is strictly true or false; therefore, probabilities are a more realistic way to view the results. The probability of being right can be represented by the *reliability* (percentage of profitable trades) of a system using these calculations. When right, the average payout is the *average profits per trade*, and when wrong the loss is the *average loss per trade*. When you can assign a probability to results, the variable can be called *fuzzy*. But this way of looking at values is really a twist on basic probability analysis.[12]

The idea of *fuzziness* is intended to describe the lack of precision in normal human conversation and thought. In its pure form, the use of fuzziness allows human uncertainty to be introduced into methods of artificial intelligence. For example, we often say:

"There were a lot of people in line at the show."
"I had to wait a long time."
"It was really cold while I was waiting."
"The stock market was strong yesterday."
"Unemployment dropped sharply."

In all these cases we understand what is being said although there are no specific values associated with "a lot of people," "a long time," "cold," "strong," and "sharply." In fuzzy logic, all is not *true* or *false*, 0 or 1, *there* or *not there*. True fuzzy logic will answer the question "If a half-eaten apple is still an apple, how much do you have to eat before it stops being an apple?"

---

[12]Parts of this section are drawn from Perry Kaufman, *Smarter Trading* (New York: McGraw-Hill, 1995).

The concept of fuzziness includes *fuzzy numbers*, such as "small," "about 8," "close to 5," and "much larger than 10," as well as *fuzzy quantifiers*, such as "almost," "several," and "most." Phrases such as "unexpected results of government reports cause big moves" is a common fuzzy expression.

Fuzzy Reasoning

Fuzzy events and fuzzy statistics are combined into *fuzzy reasoning*. It is a remarkable phenomenon that the answers to the following examples are clear to the human brain, but not to a machine (the answers can be found in the footnote):[13]

Example 1:   X is a small price move.
             Y is much smaller than X.
             How small is Y?
Example 2:   Most price moves are small.
             Most small price moves are up.
             How many price moves are up?
Example 3:   It is not quite true that the quarterly earnings were very bad.
             It is not true that the quarterly earnings were good.
             How bad were the quarterly earnings?

Common Approach to Fuzzy Solutions

In the terminology of fuzzy logic, there are three aspects of problem solving: *membership*, to show how the data are related to each other; *fuzzy rules*, to draw conclusions; and *defuzzifiers*, to turn the fuzzy answers back into usable results.[14]

For example, we would like to predict whether a price move will be big enough to capture the profit needed by our trading system. To have a robust system, we always use a 20-day trend but would like to predict which price moves will be above our minimum goal before we enter the trade. In general, we have found that, given commissions and slippage, we need to net $500 per trade. Because we are using a trend-following system that gives up profits before exiting, we would like to see an $800 peak profit in order to capture $500. Based on our needs, we can define our limits as:

Average peak move between $300 and $800 is acceptable.
Average peak move over $800 is desirable.
Average peak move under $300 is unacceptable.

---

[13]The answers to the examples are: (1) very small, (2) most, and (3) bad.
[14]Murray A. Ruggiero, Jr., "Artificial Trader Jumps Candlesticks," *Futures* (February 1995), and "Lighting Candlesticks with Fuzzy Logic," *Futures* (March 1996).

To decide whether current price movement is likely to give us enough profit to enter a trade, we measure the average price move over every 20-day period for the past 1 year (long term) and for the past 1 month (short term). Based on this approach, we create the following rules:

1. If both the long-term and short-term profit potentials are unacceptable, then the current profit potential is unacceptable.

2. If both the long-term and short-term profit potentials are desirable, then the current profit potential is desirable.

3. If both the long-term and short-term profit potentials are acceptable, then the current profit potential is acceptable.

These three cases are very clear, but what if the long-term and short-term expectations are different? For example, if the long-term average peak profits are $250 and those for the short-term are $600, what can be expected from the current move? In the current application of fuzzy logic, this converts to a problem in probability for which we need more information about the history of these price moves. Through testing, we find that the 1-year average price move has a standard deviation of $75 and the past 1 month has a standard deviation of $200. We could then construct a diagram that shows the expected results from combining the two measurements (see Figure 20.14). The average returns are shown with descending lines based on their standard deviations. These cross at about $350 per trade, giving a reasonable answer to how to combine the two values.

One additional measurement that cannot be overlooked is the potential error in each statistic. The longer-term measure used 250 trading days while the short-term used only about 20 days. The standard error for those calculations are 6.3% and 22.3%, respectively. In the final interpretation of results, we can say that the longer-term data is more
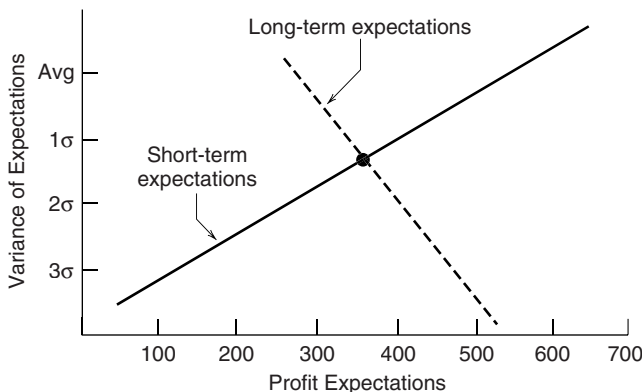
FIGURE 20.14 Expectations of a price move using two time periods.

important than the short-term values by about 3.5:1. Then we could write additional rules:

4. If the long-term potential is unacceptable and the short-term is acceptable, then the current potential is unacceptable.

5. If the long-term potential is unacceptable and the short-term is desirable, then the current potential is acceptable.

Using these rules, we can claim that we are applying fuzzy logic to our trading; however, the state-of-the-art indicates that fuzzy logic is still hanging onto the coattails of simple probability. When it lets go, it is likely to be the most significant breakthrough in market analysis.

A Candlestick Example

Ruggiero[15] has created an example of fuzzy logic to identify the bullish engulfing, evening star, and dark cloud formations of candlestick charting. These patterns can be found in Chapter 3, and are defined as

Bullish engulfing pattern. A bullish outside day follows a down day, seen as a long white body larger than the entire previous day's dark formation.

Bearish evening star. Similar to an island reversal, it is a long white body followed by a gap higher white body, followed by a gap lower dark body.

Bearish dark cloud. A long dark body followed by a lower long white body, where the second day opens above the previous day's high but closes below the midpoint of the open-to-close range.

The process also requires the following program functions (which are not given here):

```
candlecolor(open,close)
engulfing(open,close,refbar)
windowdown(open,high,low/close/lookback)
windowup(open,high,low/close/lookback)
```

The function *candlecolor* returns a 1 if white, −1 if dark. The *fuzzylong* and *fuzzysmall* functions confirm those formations, and the *windowup* and *windowdown* functions return a number greater than zero when there is a gap in the proper direction. The *lookback* parameter is the period used to calculate a moving average of the candle body size; the current candle body is considered large or small relative to the average body size. *Trigger1* is a factor that determines the current body is long when the *current*

---

[15]Murray A. Ruggiero, Jr., *Cybernetic Trading Systems* (New York: John Wiley & Sons, 1997).

body > average body × trigger1. Similarly, the body is small when the *current body < average body × trigger0*. When using the *fuzzylong* function, *trigger1 > trigger0*. The last parameter, *refbar*, identifies which bar is the one that is being engulfed.

The four key programs that identify these patterns are available on the Companion Website beginning with *TSM* to distinguish them from intrinsic functions. The logic in the programs is as follows:

FuzzyLong (function), which returns the value 1 if the candlestick is Long White and 0 if the candlestick is Long Black.

1. Calculate the range of the candle as $R_t = O_t - C_t$.
2. Find the level that represents a zero, $ztrig = average(R,n) \times zerocof$, where n is the calculation period, and n and *zerocof* are inputs.
3. Find the level that represents a one, $onetrig = average(R,n) \times onecof$, where n is the calculation period, and n and *onecof* are inputs.
4. Find the scale of the bar limits, $scale = onetrig - ztrig$. If $scale = 0$, then set scale to a large number, 99.99.
5. Define a *tall* candle as $tall_t = \max (0, \min (R_t - onetrig_t)/scale_t)$.
6. If $tall_{t-1} = 1$ and $R_{t-1} - ztrig_t <> 0$ then $tall_t = \max (1, (R_t - R_{t-1})/(R_{t-1} - ztrig_t))$
7. $Fuzzytall = tall_t$

FuzzyBullishEngulf (function), which returns 0 if it is not a Bullish Engulfing Pattern and *size* if it is. The code for *FuzzySmall* is not provided; however, it will be very similar to *FuzzyLong*. The code for *Engulfing* is not provided but should be simple to create.

1. $Sbody_t =$ FuzzySmall $(O_t, C_t, n, onecof*0.3, zerocof*1.0)$
2. $Lbody_t =$ FuzzyLong $(O_t, C_t, n, onecof*2.0, zerocof*1.0)$
3. If Englulfing $(O_t, C_t, 1) = 1$ and $color_t = 1$ and $color_{t-1}$ then

$$FuzzyBullishEngulf = \min(Sbody_{t-1}, lbody_t)$$

FuzzyEveningStar (function), which returns 0 if it is not an Evening Star and the *size* if it is.

1. $Color_t =$ candlecolor(O_t, C_t)
2. $Sbody_t =$ FuzzySmall$(O_t, C_t, n, onecof*0.3, zerocof*1.0)$
3. $FuzzyRange_t = C_t - (C_{t-2} + O_{t-2})/2$
4. If $color_t = -1$ and $color_{t-2}$ and WindowUp$(O_t, H_t, L_t, C_t, 1) > 0$ and $O_t > O_{t-1}$ and $FuzzyRange_t < 0$ then

$$FuzzyEveningStar = \min(Sbody_{t-1}, 1 - Sbody_{t-2})$$

FuzzyDarkCloud (function), which returns 0 if not a Dark Cloud and *size* if it is. In this function the inputs *onecof* and *zerocof* are reversed to test for "not small."

1. $Color_t$ = candlecolor(O_t, C_t)

2. $Sbody_t$ = FuzzySmall($O_t, C_t, n, zerocof$*0.3, $onecof$*1.0)

3. $FuzzyRange_t = C_t − (C_{t−1} + O_{t−1})/2$

4. If $color_t = −1$ and $color_{t−1}$ and $O_t > H_{t−1}$ and $FuzzyRange_t < 0$ then

$$\text{FuzzyDarkCloud} = 1 − Sbody_{t−1}$$

This is a clever and useful piece of programming that satisfies our own definition of these candlestick formations. Although it is called *fuzzy,* the logic still relies on threshold values, *trigger1* and *trigger0,* to determine that a body is long or short. It may be *dynamic,* or *adaptive,* but it may not be truly *fuzzy.*

FRACTALS, CHAOS, AND ENTROPY

Another area that has captured the interest of market analysts is chaos. *Chaos theory* is a way to describe the complex behavior of nonlinear systems, those that cannot be described by a straight line. It is also called *nonlinear dynamics.* Chaotic systems appear, at first, to be random but turn out to be "not without any form or method." One explanation is that small occurrences significantly affect the outcomes of seemingly unrelated events. If we think of the effects of the continuous flow of news on prices, the most recent may be most significant, but that fades at the same time other news arrives.

One method of measuring chaotic systems is with various geometric shapes. This effort has resulted in an area of mathematics now called *fractal geometry;* its approach strikes a true note about how the real world of numbers actually works.[16] All of us have been taught Euclidean geometry in school; it is the world of straight lines and clean edges in which we can measure the length of a line or the area of a rectangle very easily. In the real world, however, there are no straight lines; if you look closely enough—using a microscope if necessary—all "straight lines" have ragged edges and all may be described as chaotic.

Fractal Dimension

In fractal geometry we find that there is a way of representing the irregularity of numbers and the formations seen in nature. We first must accept the notion that there are no *whole* numbers in nature, that real-world objects are more likely to be described as fractional,

---

[16] An excellent discussion of this topic can be found in Edgar Peters, *Chaos and Order in the Capital Markets* (New York: John Wiley & Sons, 1991). The coastline example is originally credited to Benoit Mandelbrot. Other books can be found by searching the Internet for "Chaos Theory in Financial Markets."

or having a *fractal dimension.* The classic example of this is the algebra of coastline dimension. We will see that the questions "How long is the coastline?" and "How far did prices move?" are very similar.

The answer to both of these questions is "That depends on how it is measured." Expanding on an example earlier in this book, consider the problem of measuring the coastline of Australia using a large wall map. If we take a 12-inch ruler, and placed each end on a part of the coastline, we might find that the coastline is about 10 feet (perhaps 10,000 miles according to the scale). Using a 12-inch ruler requires that we cross over parts of the coast that jut out into the water; in other cases we might span a large harbor in order to get both ends of the ruler to touch the coastline. Had we taken a slightly smaller ruler we would have been more accurate and perhaps have found the coastline to be 15,000 miles; an even smaller ruler would have followed the contours better and found 20,000 miles of coast. As the ruler gets smaller the coastline appears to get longer. If we had an infinitely small ruler, the coast would be infinitely long. There is really no correct answer to the question, "How long is the coastline?" It all depends on your ruler.

Fractal dimension is the degree of roughness or irregularity of a structure or system. In many chaotic systems there is a constant fractal dimension; that is, the interval used for measuring will have a predictable impact on the resulting values in a manner similar to a normal distribution. Therefore, if we used a 12-inch ruler to get 10,000 miles of coastline, we might expect a predictable relationship of coastline to ruler:

24-inch ruler	5,000-mile coastline
12-inch ruler	10,000-mile coastline
6-inch ruler	20,000-mile coastline

Note that the large 24-inch ruler returns a value that is actually smaller than what we believe is a reasonable answer. This is because, when you place a long ruler from one coastal point to another on the map, it cuts across part of the land mass.

Using Fractal Efficiency

In Chapters 1 and 17, there is a discussion of Kaufman's *Efficiency Ratio*. This ratio is formed by dividing the absolute value of the net change in price movement over n periods by the sum of all component moves, taken as positive numbers, over the same n periods. If the ratio approaches the value 1.0 then the movement is smooth (not chaotic); if the ratio approaches 0, then there is great inefficiency, chaos, or noise. This same measurement has been renamed *fractal efficiency*. In Chapter 17, Kaufman related this to trending and nontrending patterns, when the ratio approached 1.0 and 0, respectively. While each market has its unique underlying level of noise, the measurement of fractal efficiency should be consistent over all markets. Markets may vary in volatility although their chaotic behavior is technically the same; therefore, theoretically, the characteristics of one market may be compared to others by matching both the fractal efficiency and volatility. In reality, there are small differences in fractal efficiency due to the

various participation and liquidity of markets; however, once these differences have been considered, the treatment of price behavior can be the same.

The interpretation of fractal efficiency as noise allows trading rules to be developed. For example, a market with less noise should be entered quickly using a trending system, while it would be best to wait for a better price if the market has been rated as high noise. A noisy market is one that continues to change direction, while an efficient market is smooth. When viewed in the long term, the level of market noise should determine the type of strategy that should be applied to each market. These characteristics are important in choosing trading rules and in turning a theoretical model into a profitable trading system.

Chaotic Patterns and Market Behavior

Chaotic patterns in prices are easy to visualize but very difficult to measure. There would be no problem in predicting price direction if every participant reacted in the same way to the same event, much the way a single planet would smoothly orbit a single sun. In the real world nothing is quite as simple. Consider the pattern of prices represented as nearby moons, *a* and *b*, that are each circling nearby planets of equal mass, *P1* and *P2*. Because the planets are larger than the moons, they are called *attractors*. We will get a wobbly pattern whenever the moon passes across the midpoint where one attractor is stronger than the other, shown as a straight line in Figure 20.15a. A moon at point *a* is
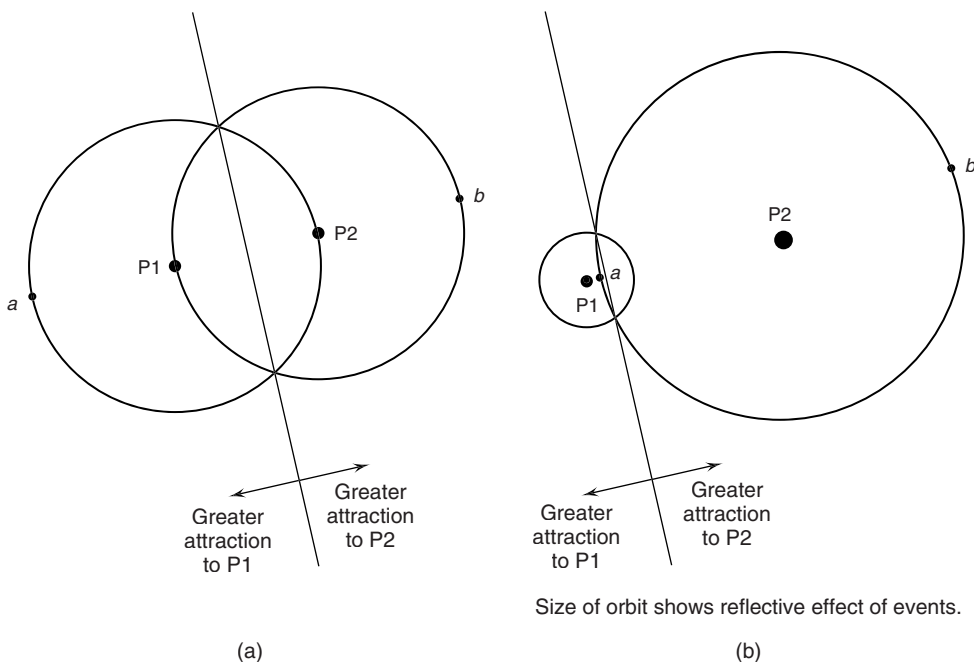
FIGURE 20.15   (a) Equal attractors cause a symmetric pattern, often a figure 8, that switches between P1 and P2. (b) Very unequal attractors show only a small disturbance in a regular orbit.

most affected by the nearest attractor, *P1*, but as it circles it becomes closer to *P2* and tries to form an orbit around that planet. It might form a figure-8 pattern as it switches between *P1* and *P2*, ending up in location *b* on the far side of *P2*. The possible patterns are too complex, and they vary based on the distance between *P1* and *P2* and the size of *P2* compared to *P1*. If attractor *P2* is much larger than *P1*, as in Figure 20.15b, there will simply be a distortion in the orbit around *P2* at the location of the straight line, rather than a switching of orbits from one planet to another.

As complex as these patterns might get, they are simple when compared to reality. Each day brings events of varying importance into the market, acting as attractors. Each attractor has an initial importance that loses value over time. To make matters worse, we cannot always predict when a new attractor, or news event, will appear. When we can predict the time, such as an economic report, we cannot predict the impact. This makes the chaotic pattern very similar to raindrops falling on a pond. Each new drop, equivalent to a new event, hits at an unpredictable time and place, with variable size, and forms circular ripples. These ripples dissipate as they get further away from the point of contact in the same way that the importance of an event fades away over time. The interesting aspect of the raindrop analogy is that, while we cannot predict where the next raindrop will fall, once it has landed we can completely determine its effects—until the next drop hits. This analogy is remarkably similar to the market. The frequency and size of price shocks are like rain drops; occasionally there is a large one that overwhelms the smaller events, the market noise, for a short time.

Entropy—Predicting by Similar Situations

Forecasting the future by finding a similar combination of factors, *similar conditions*, that occurred in the past has always been a technique used by fundamental analysts. It is necessary to match the state of major economic indicators with the less quantifiable government policy, then look at how the prices reacted to these similar setups.[17] The entire area of charting, traditionally called *technical analysis*, is based on a predictable outcome following a specific chart pattern. These patterns can be simple or very complex, exist in a single time frame or be the composite of multiple time frames.

There is one mathematical method that comes closest to being able to identify similarities in price patterns. *Conditional entropy* can measure the probability that two patterns are similar. The higher the value of conditional entropy, the more likely the results of a past pattern will predict the results of a current pattern.[18] This has an analogy in *autocorrelation*, which measures the ability of past data to predict future data.

---

[17]A data service, LIM (Logical Information Machines), provides both extensive fundamental and price data along with tools for comparing past events and their outcomes.

[18]The mathematics for conditional entropy, with excellent examples, is located on the website http://tecfa.unige.ch and was written by Phillippe Lemay. This section is adapted from his work.

Entropy

Entropy, the basis for this more advanced form, is a measurement of expectations. If the output of a system is known, and there is no uncertainty, the value of the entropy formula is zero. When entropy is low, the state of the system is easy to predict. As the output becomes less certain, the value approaches 1, and there is more disorder. Entropy, H, is defined as a function of price, x,

$$H(x) = -\sum_{x=i}^{n} p_i(x)\log_2 p_i(x)$$

where x can have n different values, and $p_i(x)$ is the probability that the outcome will be the ith value of x. When the system has only one possible outcome, $n = 1$, then $H(x) = 0$. If all n possible outcomes have an equal probability, then uncertainty is maximized, and $H(x) = \log(n)$. Therefore, if there are 16 possible outcomes, the maximum entropy is 4. It is convenient to normalize the value of the entry by dividing by the maximum value $\log(n)$.

As an example, consider the price moves that follow a particular chart pattern, where prices close in the upper 10% of the daily trading range and the stock is above its 200-day moving average. The hypothetical frequency of the next day's move is shown in Table 20.4. The calculations in the columns follow the formula $H(x)$ and show the result $H(pattern) = 2.20$. The maximum entropy for a 6-scale variable is $\log_2(6) = 2.58$; therefore, there is considerable uncertainty in the ability of this pattern to forecast 1 day ahead.

Conditional Entropy

Conditional entropy will give us the probability of the outcome based on the current price pattern being similar to a previous price pattern. But first we need to create a frequency distribution to use as a basis for the calculations. Table 20.5 gives the number of times that an upwards price change occurred given the completeness of a chart pattern while the current price was above the 200-day moving average. The value 1 on the left scale indicates that the pattern was complete; the value 6 is a minimum recognition of the pattern. The top scale shows the percentage change in price on the next day. The

TABLE 20.4 Frequency of Price Moves Following a Known Pattern, plus Entropy Calculations

Price Move in %	Frequency	Probability	log2(prob)	prob × log2(prob)
+3	33	0.13	−2.98	−0.38
+2	87	0.33	−1.58	−0.53
+1	84	0.32	−1.64	−0.53
−1	35	0.13	−2.90	−0.39
−2	14	0.05	−4.22	−0.23
−3	8	0.03	−5.03	−0.15
Sum	261	1.00	−18.35	−2.20

TABLE 20.5 Frequency of Price Moves Following the Completeness of the Chart Pattern When Prices Are above the 200-Day Moving Average. The value 1 represents the completed pattern and 5 is the beginning of the pattern.

Completeness of Chart Pattern	Price Change Next Day in %					
	+3	+2	+1	−1	−2	−3
1	19	4				
2	11	63	64	3	1	
3	2	16	18	20	2	2
4	1	4	1	9	6	2
5			1	2	4	3
6				1	1	1

frequencies show that price changes of 1% and 2% were most common when the chart pattern was near completion but not yet fully completed.

Conditional probabilities can be calculated from the frequency distribution in Table 20.5. These are shown in Table 20.6. To compute the conditional probability of state j of variable Y (the dependent variable) given state i of variable X (the independent variable, the past pattern), we write $p(Y = j | X = i)$. Then the conditional probability is computed by dividing the frequency of occurrence of the two states, n_{ij}, by the total frequency of state i, given as n_i. Then the probability of Y conditioned on X is

$$p_{j|i} = \frac{n_{ij}}{n_i}$$

In this hypothetical example, the probability is highest, and the price moves the most, when the chart pattern is complete. In a test with real data, we might expect the two values in row 2, columns +2 and +1 to show the greatest probabilities. The values in the lower right corner are high, but are based on a sample of only 1 value; therefore, they are not reliable.

The conditional probabilities can then be entered into the final formula for conditional entropy. The conditional entropy is the average sum of the entropies based on the

TABLE 20.6 Conditional Probabilities of a Price Change Given the Completeness of a Pattern. Values correspond to Table 20.5.

Completeness of Chart Pattern	Price Change Next Day in %					
	+3	+2	+1	−1	−2	−3
1	0.83	0.17				
2	0.08	0.44	0.45	0.02	0.01	
3	0.03	0.27	0.3	0.33	0.03	0.03
4	0.04	0.17	0.04	0.39	0.26	0.09
5			0.1	0.2	0.4	0.3
6				0.33	0.33	0.33

conditional probability, *H*, of *Y* (the second event or price change) conditioned on *X* (the first event or completeness of pattern)

$$H(y\,|\,x) = -\sum_{x=i}^{t_1} p_i(x) \sum_{y=j}^{t_2} p_{ij}(y\,|\,x) \log_2 p_{ij}(y\,|\,x)$$

The higher the value of *H*, the greater the predictive value.

NEURAL NETWORKS

Neural networks are recognized as a powerful tool for uncovering market relationships.[19] The technique offers exceptional ability for discovering nonlinear relationships between any combination of fundamental information, technical indicators, and price data. Its disadvantage is that it is so powerful that, without proper control, it will find relationships that exist only by chance.[20]

Although the idea and words for the computerized neural network are based on the biological functions of the human brain, an *artificial neural network* (ANN) is not a model of a brain, nor does it learn in the human sense. It is simply very good at finding patterns, whether they are continuous or discrete, appearing at different times.

The operation of an artificial neural network can be thought of as *a feedback process*, similar to the Pavlovian approach to training a dog:

1. A bell rings.
2. The dog runs to 1 of 3 bowls.
3. If right, the dog gets a treat; if wrong, the dog gets a shock.
4. If trained, stop; if not trained, go back to Step 1.

Terminology of Neural Networks

The terminology used in the computerized neural network is drawn from the human biological counterpart, shown in Figure 20.16. The following are the principal elements.

- *Neurons* are the cells that compose the brain; they process and store information.
- *Networks* are groups of neurons.

---

[19] Parts of this section are based on Perry Kaufman, *Smarter Trading* (New York: McGraw-Hill, 1995), 164–171.

[20] Readers are referred to E. Michael Azoff, *Neural Network Time Series Forecasting of Financial Markets* (New York: John Wiley & Sons, 1994); Robert Trippi, *Neural Networks in Finance and Investing*, 2nd ed. (New York: McGraw-Hill, 1996); Apostolos-Paul Refenes, *Neural Networks in Capital Markets* (New York: John Wiley & Sons, 1995); and Edward Gately, *Neural Networks for Financial Forecasting* (New York: John Wiley & Sons, 1996), for a more introductory approach. It seems that very little has been published during the past 10 years although there is more software available.
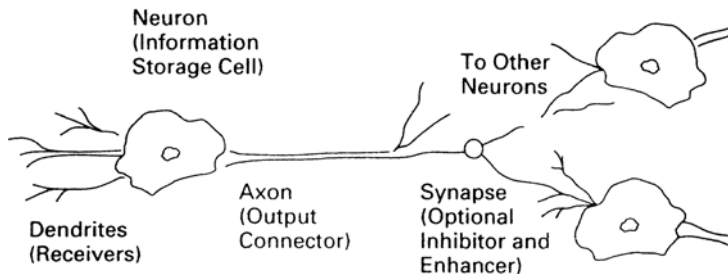
FIGURE 20.16   A biological neural network. Information is received through dendrites and passed to a neuron for storage. Data are shared by other cells by moving through the output connector, called an axon. A synapse may be located on the path between some individual neurons or neural networks; they select the relevant data by inhibiting or enhancing the flow.

Source: Perry Kaufman, *Smarter Trading* (New York: McGraw-Hill, 1995), 165.

- *Dendrites* are receivers of information, passing it directly to the neurons.
- *Axons* are pathways that come out of the neuron and allow information to pass from one neuron to another.
- *Synapses* exist on the path between neurons and may inhibit or enhance the flow of information between neurons. They can be considered *selectors*.

Artificial Neural Networks

Using essentially the same structure as a biological neural network, the computerized or artificial neural network (ANN) can generate a decision on the direction of the stock market.[21] It relies heavily on the synapses, which are interpreted as weighting factors in this process. To achieve its result it will also combine inputs that interact with one another into a single, more complex piece of data using layers of neurons, as shown in Figure 20.17, a classic 3-layer neural network. The ANN will modify these weighting factors in order to minimize the output error when compared to known results (the learning phase). It differs from a least squares regression because it includes linear, nonlinear, and pattern recognition relationships.

To be most efficient and avoid overfitting, the inputs to the ANN should be those factors considered most relevant to the direction of stock prices. The five items chosen here were the Gross Domestic Product (GDP), unemployment, manufacturing or inventories,

---

[21]The website www.cs.stir.ac.uk, featuring a description by Professor Leslie Smith, Centre for Cognitive and Computational Neuroscience, University of Stirling, U.K., was found to have very useful information.
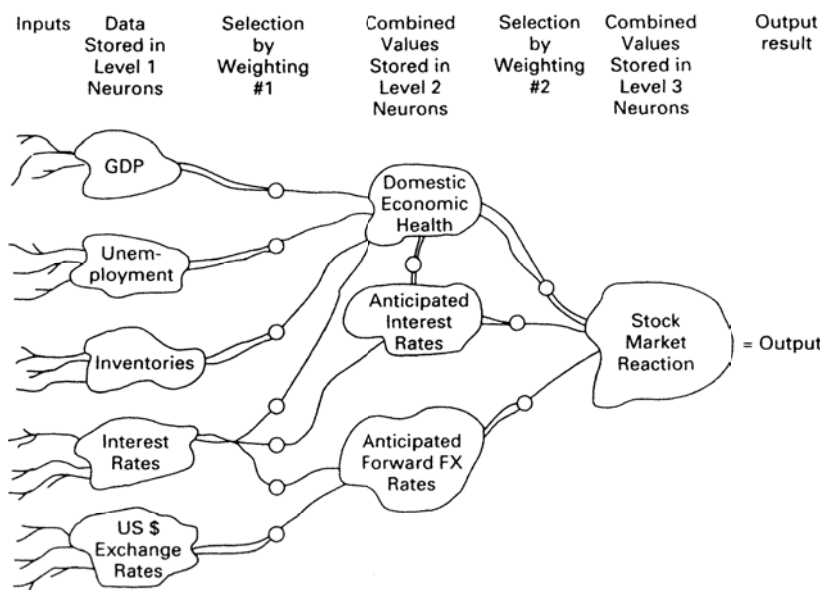
FIGURE 20.17 A 3-layer artificial neural network to determine the direction of stock prices.

Source: Perry Kaufman, *Smarter Trading* (New York: McGraw-Hill, 1995), 166.

interest rates, and the value of the U.S. dollar—all readily available data. Each of these items is input and stored in separate Layer-1 neurons. Changing the values in these neurons may have a positive or negative effect on the final output, which is the expected direction of stock prices. An improved GDP, lower unemployment, lower inventories, and a lower U.S. dollar are all good signs for economic growth and result in the possibility of higher interest rates. Higher rates are a defensive action by the Federal Reserve to prevent inflation with the eventual consequence of a decline in stock prices. Interest rates themselves have a direct effect on stock prices, improving corporate profits as the cost of borrowing declines.

Each neuron in Layer 1, which receives the data from the dendrites, is connected to a second layer of neurons through a synapse. Each synapse can be used to restrict or enhance information by assigning a weighting factor. For example, if changes in unemployment have a greater ultimate impact on stock prices than changes in the GDP, then it may receive a weight of 1.5 compared to 0.9 for the GDP. If very small changes in any of the data are deemed unimportant to the result, then the synapse can act as a threshold and only allow data to pass if it exceeds a minimum value. If the data is irrelevant, it gets a weighting factor of zero; if it reacts contrary to the result, the weighting factor will be negative.

The second layer of neurons is used to combine initial data into significant subgroups by trial and error. To allow this to happen, every neuron in Layer 1 must be connected to every neuron in Layer 2 by its own axon containing a synapse. Figure 20.17 shows the

significant groupings that might result, ignoring the other connecting axons which would clutter the illustration. GDP, unemployment, and inventories are combined into a single item called Domestic Economic Health (DEH). The synapses allow each element to be assigned a specific level of importance relative to DEH using weighting factors. Also note that the neuron DEH is altered by anticipated interest rates, which is the result of data flowing to another neuron in Layer 2. Finally, the three neurons in Layer 2 are combined according to importance, giving the net stock market reaction to the input data.

Layer 2 may be called a *hidden layer*, because you normally only see the input and output layers. Layer 2 has the complex interconnections that allow feedback to occur. It is possible to have more complex connections by adding another hidden layer. In general, the more hidden layers, the longer the solution time and the more finely tuned the answer.

The human brain works in a way very similar to the artificial neural network shown in Figure 20.17. It groups and weighs the data, combines them into subgroups, and finally produces a decision. The human process of weighing the data is complex and not necessarily transparent; that is, we may never know the precise flow of data, how many layers exist, and how weights are assigned and reassigned before the final decision.

Because a computerized neural network cannot *know* whether its answer is correct, you must tell it. This is done by giving the computer the historic data and the corresponding answers. By giving the artificial neural network a long history of information, it can determine, using *feedback*, the weighting factors that would have given the correct results most often. The more history that is given to the computer, the more likely it will find a robust answer. Figure 20.18 shows the results of using five different inputs to
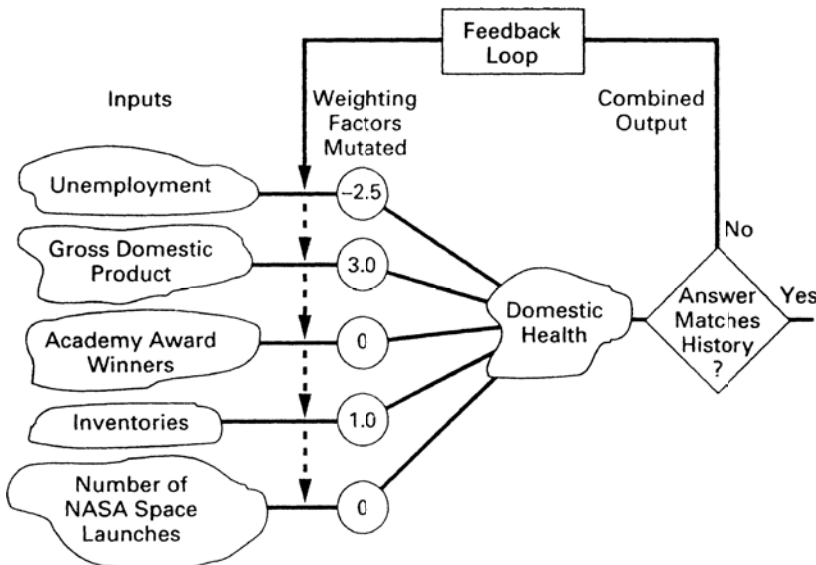
FIGURE 20.18   Learning by feedback.
Source: Perry Kaufman, *Smarter Trading* (New York: McGraw-Hill, 1995), 166.

predict the direction of the stock market. Two of the inputs, the Academy Award Winners and the Number of NASA Launches, are not likely to be useful in the long term, but *may appear* to provide valuable information about economic activity for short intervals. By using enough comparisons, the weighting factors are found to show that unemployment has a strong negative effect on prices, the GDP a strong positive effect, and inventories have a weak positive effect. The other items had no consistent predictive ability and received a weight of zero. This feedback process is called *training*.

Selecting and Preprocessing the Inputs

There is considerable debate over the inputs needed to find a successful solution using a neural network; however, everyone agrees that the selection of inputs is critical. These inputs must be presented in the most direct form because the neural network will not be able to change them. This step is called *preprocessing*. We must decide which factors are most likely affecting the direction of stocks and the ability to anticipate that direction, then prepare data that contains information with those qualities. For example, we may want to know the short-term and long-term trends, the direction of interest rates, the Dow Jones Utility Index, the ratio of interest rates to gold, economic data such as the GDP and the Balance of Trade, technical indicators such as the RSI, stochastic, and ADX, and a 20-day moving correlation between the U.S. stock market and other major equity markets. There are countless factors that might influence the direction of stocks; the more you choose, the slower the solution and the greater the chance of a less robust model. If you choose too few, they may not contain enough information; therefore, the preprocessing problem requires practice. You may also construct a number of simple trading systems that show profits and include their basic components as inputs to the neural network. You might create a performance series for a specific system that has only values –1, 0, and 1, representing short, neutral, and long market positions. In that way, the neural net may be used to enhance an existing trading strategy.

Selecting the Success Criteria

In a manner similar to evaluating any backtesting result, the training process requires that you select the success criteria (a full discussion of this can be found in Chapter 21). One possibility for the success criteria is a combination of frequency of trading, the size of the profits per trade, a reward/risk ratio, and a frequency of profitability.

According to Ruggiero,[22] neural networks can be used to predict price direction, such as the percentage change five days into the future; however, they are much better at predicting *forward-shifted technical indicators* because these tend to have smoother results. Most technical indicators, such as trend and momentum calculations, smooth the

---

[22]Murray A. Ruggiero, Jr., "Build a Real Neural Net," *Futures* (June 1996), the first of a series of excellent articles on this subject.

input data, which is most often the price. The longer the time period used in the calculation, the greater the smoothing. Ruggiero suggests an output function such as

```
Average((close[−5] − Lowest(Close[−5],5))/(Highest(close[−5],5) −
  Lowest(close[−5],5)),5)
```

which is the 5-day average of a 5-period stochastic shifted forward by 5 periods. In this notation, the bracketed value [–5] represents 5 periods into the future.

The Training Process

At the heart of the neural network approach is the feedback process used for training shown in Figure 20.18. This is the part of neural networks that many people refer to as the *learning process*. To observe sound statistical practice, it is advised that only about 70% of the data be used for training. As the neural net refines the weighting factors for each of the inputs and combinations of inputs (in the layers between the input and output), it will need more data to test these results. Of the remaining 30% of the data, 20% should be designated for testing. The success or failure of the method to find a solution is based on the performance of the test data. The remaining 10% is saved for out-of-sample validation.

Weighting factors are found using a method called a *genetic algorithm*, discussed in the next section and extensively covered in Chapter 24. For now, we need to know that the training process begins with an arbitrary or random value assigned to the weighting factors of each input. As the training proceeds, these weighting factors are randomly *mutated*, or changed, until the best combination is found. The genetic algorithm changes and combines weighting factors in a manner referred to as *survival of the fittest*, giving preference to the best and discarding the worst.

Testing is completed when the results, as measured by the success criteria, cannot be improved. That is, after a number of feedback loops on the historic data (the oldest 70% of the data), the test data (the next 20%) is used with the new weighting factors. If the results are improving (also called *converging*) the process continues. If the results are improving at a very slow rate, or have stopped because a better combination cannot be found, the neural network process is completed. Sometimes the results get worse rather than better. This can be fixed by beginning again using new random weighting factors. If this fails to solve the problem, then it is most likely that there are too many inputs that are irrelevant. Starting with a smaller number of basic inputs should get the process back on track.

Because the genetic algorithm is a trial-and-error process, rather than an analytic approach, the best results could be found by chance, rather than by cause and effect. With enough data, it is always possible that two series of events will appear related even when they are not. It is necessary to review the results to avoid obvious mistakes.

A Training Example

We would like to train an ANN to tell us whether we should buy or sell stocks. As inputs, we select what we believe to be the five most relevant fundamental factors: GNP,

TABLE 20.7   Two Training Cases (Initial State)

	Case 1			Case 2		
Input	**Relative Value**	**Weight**	**Net**	**Relative Value**	**Weight**	**Net**
GNP	Strong	50	1.0	50	Weak	−60	1.0	−60
Unemployment	Low	−25	1.0	−25	High	40	1.0	40
Inventories	Low	−50	1.0	−50	Neutral	15	1.0	15
U.S. dollar	Very Strong	75	1.0	75	Neutral	0	1.0	0
Interest rates	Falling	−25	1.0	−25	Rising	45	1.0	45
Total value		75				40	
Threshold		±125				±125	
Current response		None				None	
Actual market reaction		Strong				Weak	

Source: Perry Kaufman, S*marter Trading* (New York: McGraw-Hill, 1995).

unemployment, inventories, the U.S. dollar index, and short-term interest rates. This test does not use any preprocessed data, such as trends or indicators. To simplify the process, the following approach is taken:

1. Each input is normalized so that it has values between +100 and −100, indicating strength to weakness, with 0 as neutral.

2. When the combined values of the five indicators exceeds +125, we will enter a long position; when the combined value is below −125, we will enter a short.

3. Values between +125 and −125 are considered neutral to the trading strategy.

To show how the training process works, two combinations of starting values are shown in Table 20.7 and Table 20.8 as Case 1 and Case 2. Table 20.7 is the initial state

TABLE 20.8   Two Training Cases (After Mutated Weighting Factors)

	Case 1			Case 2		
Input	**Relative Value**	**Weight**	**Net**	**Relative Value**	**Weight**	**Net**
GNP	Strong	50	1.0	50	Weak	−60	1.0	−60
Unemployment	Low	−25	−1.0	25	High	40	−1.0	−40
Inventories	Low	−50	1.0	−50	Neutral	15	1.0	15
U.S. dollar	Very Strong	75	1.0	75	Neutral	0	1.0	0
Interest rates	Falling	−25	−1.0	25	Rising	45	−1.0	−45
Total value		125				−130	
Threshold		±125				±125	
Current response		None				None	
Actual market reaction		Strong				Weak	

Source: Perry Kaufman, *Smarter Trading* (New York: McGraw-Hill, 1995).

of the neural network, where we begin by arbitrarily setting all of the weighting factors to 1.0. In actual training, the network might require that the sum of all weights total 1.0. The actual values of the normalized inputs are shown in the columns marked *Relative Value*, and the correct historic answers are at the bottom, marked as *Strong* and *Weak* stock market reactions to these values. For the neural network to return the correct answers, it must produce a result greater than +125 for Case 1 and below −125 for Case 2. By assigning initial weights of 1.0 to all inputs, the value of Case 1 is +75, and the value of Case 2 is +40; both fail to produce the known historic answer and testing continues searching for better weighting factors.

In Table 20.8 the weighting factors have undergone mutation using a genetic algorithm. This example attempts to use only weighting factors of +1.0, 0, and −1.0. By reversing the effect of unemployment and interest rates on the direction of the stock market, and by random selection of weighting factors, the results now match the historic pattern of stock movement. Using fractional values for weighting factors, and many more training cases, the ANN method should find the current underlying relationship between these inputs and stock market movement.

Success Criteria

Determining success during the learning process is a matter of measuring the ANN output against the training data, looking for convergence. The common measurements are the average correlation (adjusted for the number of parameters), the *t*-statistic, the t^2-statistic, and the *F*-statistic. The t^2-statistic is unique to neural networks and measures the nonlinear relationships between two variables[23]

$$t^2 = \sqrt{AS} \times \frac{\sigma_I}{\sigma_O} \times \sqrt{n-k}$$

where σ_I is the standard deviation of the input data
σ_O is the standard deviation of the output data
n is the number of observations
k is the number of model inputs
AS is the average squared slope of the cross-section

As with most comparisons, it is also necessary to include a measurement of accuracy based on the size of the sample. The *root mean square* (*E*) and the *mean absolute percentage error* (*MAPE*) can be used,

$$E = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n}}$$

---

[23] Stephan Kudyba, "Are Neural Networks a Better Forecaster?" *Futures* (October 1998).

$$MAPE = \frac{\sum_{i=1}^{n} \frac{|e_i|}{|A_i|}}{n}$$

where n is the number of observations,
        e_i is residual error of observation i
        A_i is the actual value of observation i

Reducing the Number of Decision Levels and Neurons

The robustness of a neural network solution is directly related to the number of decision levels, the number of neurons in each level, and the total number of inputs. Fewer elements produce more generalized and therefore more robust solutions. When there are many decision layers and many neurons, the inputs can be combined and recombined in many different ways, allowing very specific patterns to be found. The more specific, the greater the chance that the final solution will be overfitted, that is, it will be fine-tuned to such specific patterns in the past that those patterns will not occur in the future.

Failure to converge to a solution can also occur if there are too many inputs and the genetic algorithm allows too much mutating (discussed in the next section). Neural networks can be highly complex and require experience before they can be used efficiently.

The trade-offs using a neural network are the same as with most other optimization methods. Too many inputs and combinations increase the time of testing and increase the chance of a solution that is overfit. Too few values can produce a result that is too general, has large risk, and is not practical. It is best to begin with the most general and proceed in clear steps toward a more specific solution. In this way you will understand the process better, ultimately save time, and be able to stop when you have reached the practical limitations of this method.

Modeling Human Behavior

Neural networks are considered a learning process similar to the parallel architecture of the human brain. It can be trained on a classification of data, then used to find new data of the same class, or it can be trained on a number of points in a function (for example, a logarithmic curve or more complicated cycle), then used to find other points on that function.

One application that is relevant to price forecasting is to make each of the Layer-1 neurons contain the results of the major economic indicators. The synapse between level 1 and level 2 would discard any indicator when its value was not significantly different from zero. This reflects the way the public views economic reports. During the scenario of economic recovery, such as early in 2003 as well as after the subprime crisis of 2008, corporate earnings and income growth were most important. Other reports

were, for the most part, ignored. After earnings improved and the stock market had rallied, traders looked for employment statistics to improve as a means for sustaining economic growth and the stock market rally. At that point, earnings were no longer as important as more jobs. A neural network can be constructed to reflect this selection process that makes one or two economic reports more important than others given the state of the economy. Using previous GDP, stock market trends, and the direction of Fed funds, we can limit the entire neural net solution to the relationship between one or two economic indicators and the future direction of stock prices, setting the synapse weighting factors to zero for the other reports in Layer-1 neurons. A similar method might have the synapses assume a weighting factor of 1 when the economic report has a change that is large enough to attract the attention of the public (at least a 1.5 standard deviation change) and a value of 0 at any other time. In that way a minor change in all indicators does not have a cumulative impact on the neural network while the public shrugs off the data.

A Final Comment about Neural Networks

When using neural networks and other powerful searching tools, every analyst must keep in mind that neural nets will nearly always find the correct solution in a *closed system*, one where the output is fully defined by all the inputs. An ANN is often used when it is too difficult to find a solution by other means. For example, in a chemical or manufacturing process, the formula may have too many variables. A neural network will not only solve the problem, but solve it quickly. The stock market, however, is different. No one is sure that there is a solution to the problem *as we are specifying it*. Because of this uncertainty, there is no guarantee that the answer will work. You will then need to test the results on out-of-sample data to verify the success.

GENETIC ALGORITHMS

The concept of a genetic algorithm is based on Darwin's theory of survival of the fittest. In the real world, a mutation with traits that improve any creature's ability to survive will continue to procreate. This has been applied to system development using a technique called a *genetic algorithm*. Although a genetic algorithm[24] is actually a sophisticated search method that replaces the standard optimization, it uses a technique that parallels the survival of the fittest. It is particularly valuable when the number of tests or combinations is so large that a test of all combinations is impractical. Instead of the typical sequential search, it is a process of random seeding, selection, and combination to find the best set of trading rules or parameter values. Standard statistical criteria are used in the selection process

---

[24]This technique is attributed to John H. Holland, *Adaptations in Natural Language and Artificial Systems* (Ann Arbor: University of Michigan Press, 1975).

to qualify the results. Searching for a large, optimal set of parameters or finding the best portfolio allocation (as covered in Chapter 24) takes minutes using a genetic algorithm; a standard sequential search may take weeks at the same computing speed.

Representation of a Genetic Algorithm

Using the words common to this methodology, the most basic component of a genetic algorithm is a *gene;* a number of genes will comprise an *individual*, and a combination of individuals (and therefore genes) is a *chromosome*. Each chromosome represents a potential solution, a set of trading rules or parameters where the genes are the specific values and calculations. These in turn form individuals that represent rules that ultimately form a trading strategy. For example, Chromosome 1 might be a rule to buy on strength:

> If a 10-day moving average is less than yesterday's close and a 5-day stochastic is greater than 50, then *buy*.

Chromosome 2 could be a rule that buys on weakness:

> If a 20-day exponential is less than yesterday's low and a 10-day RSI is less than 50, then *buy*.

If we rewrite these two chromosomes in a notational form, the genes and individuals in its structure become more apparent:

$$\text{Chromosome 1: MA, 10, <, C, [0], \&, Stoch, 5, >, 50, 1}$$
$$\text{Chromosome 2: Exp, 20, <, L, [1], \&, RSI, 10, <, 50, 1}$$

Each of these chromosomes has 11 genes, any one of which can be changed. In addition, each chromosome has two individuals, separated by an "&" operator. In Table 20.9, the description of the genes indicates other values that can replace the current ones.

TABLE 20.9 Functional Description of the Genes in Chromosomes 1 and 2

Gene	Chr 1	Chr 2	Function
1	MA	Exp	Trend type (moving average, exponential smoothing, breakout)
2	10	20	Calculation period for gene 1
3	<	<	Relational operator (<, <=, >)
4	C	L	Price used in gene 1 calculation ((H + L + C)/3, indexed value)
5	0	1	Reference to past data ([0] = current day, [1] = previous day)
6	&	&	Method of combining individuals (and, or)
7	Stoch	RSI	Indicator type (RSI, stochastic, MACD)
8	5	10	Calculation period for gene 7
9	>	<	Relational operator (<, <=, >)
10	50	50	Comparison value for relational operator
11	1	1	Market action (1 = Buy, −1 = Sell)

Table 20.9 is a way to represent the chromosomes and individuals in a general form. It is easy to see that each gene can be changed, and each change will represent a new trading rule. For example, the 10- and 20-day averages in gene 2 of chromosomes 1 and 2 could be changed to 5 and 15 days; or, the indicators *Stoch* and *RSI* could be changed to *MACD* and *Momentum.* A combination of trading rules, or chromosomes, will create a trading strategy. Before continuing, the following steps will be needed to use the genetic algorithm to find the best results:

1. A clear way of representing the chromosomes and their component individuals and genes.

2. A *fitness* criterion to decide that one chromosome is better than another.

3. A *propagation* procedure that determines which chromosomes will survive and in what manner.

4. A process for *mutation* (introducing new characteristics) and *mating* (combining genes) to give chromosomes with greater potential a better chance for survival.

Representation of the genes and chromosomes has been shown in Table 20.9. However, there is one additional step needed to get started: creating an initial chromosome pool.

Initial Chromosome Pool

To begin, an initial pool of chromosomes must be created. These will look just like those in Table 20.9; however, each gene will be chosen at random from a list of candidates that has been previously defined. Eleven lists are needed, one for each unique gene.

1. *Trend type,* 1 of 5 choices: a moving average, exponential smoothing, linear regression, breakout, or step-weighted average.

2. *Trend calculation period,* a number between 1 and 200.

3. *Trend relational operator,* 1 of 3 choices: <, <=, >.

4. *Price used in the trend calculation,* 1 of 4 choices: C, (H + L + C)/3, (H + L)/2, indexed value.

5. *Reference data or lag:* a number between 1 and 10.

6. *Method of combining individuals,* 1 of 2 choices: *and* or *or.*

7. *Indicator type,* 1 of 5 choices: RSI, stochastic, MACD, momentum, Fisher transform (all indicators must be transformed to return values between −100 and +100).

8. *Indicator calculation period,* a number between 1 and 50.

9. *Relational operator,* 1 of 2 choices: > or <.

10. *Comparison value for indicator and relational operator,* a number between −100 and +100.

11. *Market action,* 1 of 2 choices: *buy* or *sell.*

These 11 lists offer the possibility of 4.8×10^{10} combinations, or 48,000,000,000 possibilities. Regardless of the computer speed, trying each combination would be a daunting task. However, the genetic algorithm is likely to find the best solution, or a solution very nearly the best, in less than 1,000 passes through its process.

To assure a fast solution, it is necessary to start with a large chromosome pool. Each chromosome in the pool is created by selecting one item from each list randomly. This pool represents the *genetic population*.

Fitness

Having created a pool of chromosomes that look similar to those in Table 20.9, it is necessary to decide which are best. This is done by defining a *fitness* criterion, or *objective function*, which can be used to rank the chromosomes. Because fitness will lead to survival, it is very important to decide which chromosomes should be discarded and which should survive, or even flourish. A fitness criterion must combine the most important features associated with a successful trading strategy:

- Net profits or profits per trade.
- The number of trades or a sample error criterion.
- The smoothness of the results or a reward-to-risk ratio

Ideally, preference is given to systems that have large profits, lots of trades, and very consistent performance. To measure that result, the following might be used:

$$Rank = PPT \times \left(1 - \frac{1}{\sqrt{NT}}\right) \times \frac{GP}{GL}$$

where PPT = the profits per trade
            NT = the number of trades
            GP = the gross profits
            GL = the gross losses

Alternatively, the term GP/GL can be replaced by the information ratio, AROR/ASD, the annualized rate of return divided by the annualized standard deviation, or volatility.

The fitness criterion *rank* favors large profits per trade, a large number of trades, and a high ratio of gross profits to gross losses, which is a measure of smoothness of returns. These factors are not weighted, but they serve as a simple measurement for ranking the chromosomes. For example, a trading method that returned $500 per trade for 10 trades, with a gross profit/loss ratio of 0.5 would have a rank of 170.75. Another system that returned only $250 per trade over 100 trades with a profit/loss ratio of 1.4 would rank 315.00. Therefore the system with the smaller returns has a much more agreeable trading profile according to the fitness criteria based primarily on the gross profit to gross loss ratio. Each analyst must create a criterion that allows those strategies that are most desirable to survive.

Propagation

The process of *natural section* allows only the best individuals to survive. A strong *propagation* criterion is used to encourage the survival of chromosomes with the highest ranking, as determined by the fitness test. A weak criterion allows lower rankings to survive. The implementation of this selection process is also drawn from evolution. When an individual has a high fitness score, it is allowed to create more offspring; therefore, it becomes a larger part of the population. When it has a low score, it creates fewer offspring, or no offspring, and eventually disappears from the population.

The propagation step in a genetic algorithm determines how many offspring will be created from each existing chromosome based on its fitness score. For example, if there are only 10 chromosomes in the initial pool, with fitness scores from 1 to 5 (5 being the best), they are sorted in descending order of fitness as follows

Chromosome rank	1	2	3	4	5	6	7	8	9	10
Fitness score	5	4	3	3	2	2	2	1	1	1

The propagation criterion states that the fitness score of each chromosome is multiplied by 3 and divided by 5 in order to get the number of new offspring when the chromosome is propagated. Then the numbers of offspring, in the same order, are

Chromosome rank	1	2	3	4	5	6	7	8	9	10
New offspring	3	2.4	1.8	1.8	1.2	1.2	1.2	.6	.6	.6

The fractional number of offspring can be rounded or truncated (the fractional part dropped off) to get whole numbers. If rounded, there will be 3 offspring for the highest ranking chromosome, 2 for the next, 2 for the next 2, and 1 for the next 6, for a total population of 15. But it is necessary to keep the population fixed at 10 because the pool size should represent the maximum space available in the computer;[25] therefore, the bottom 5 chromosomes do not survive. The new population, using the original chromosome ranking is 1, 1, 1, 2, 2, 3, 3, 4, 4, and 5. There are more offspring from the fittest chromosome 1; therefore, it has a better chance of surviving. The next few in line have a reasonably good chance of survival, and the ones at the bottom have disappeared.

One popular way to calculate the number of offspring is to divide the fitness of one chromosome, *f*, by the average fitness of all chromosomes, *F*. Therefore,

$$Number\ of\ offspring = Current\ fitness/Average\ fitness = f/F$$

This method works best when there are a large number of choices rather than only 10 possibilities. However, if the fitness of the linear regression chromosome was 4, and

---

[25]The need to create a pool of maximum size, and maintain that size, is discussed in Chapter 24.

the average of all chromosomes was 2, the linear regression would propagate 2 or 3 off-spring. The more offspring, the better chance of survival during both the real evolution-ary process as well as the genetic algorithm process.

Mating

The new pool of offspring now moves to the next Darwinian test, *mating*. Remember-ing that a genetic algorithm is a sophisticated search method, it needs a way to combine the genes of individuals that have passed the fitness criteria in different ways so that we might create a better chromosome. This is very similar to cross-pollination, where you continue to pollinate two species with certain desirable but different features, hoping to get a new species that has the best of both. This is done through *mating* (also called *crossover*).

When mating, we must be careful to combine, or switch, the entire individual, and not just one gene (remember that individuals are separated by an "&" symbol). All of the genes that form an individual make sense only as a set. If we separate the individual, then the result could be an inferior offspring. Using chromosomes 1 and 2, we can switch the first individuals to get

<div align="center">

Offspring 1: Exp, 20, <, L, [1], &, Stoch, 5, >, 50, 1

Offspring 2: MA, 10, <, C, [0], &, RSI, 10, <, 50, 1

</div>

In the actual genetic algorithm process, only a small percentage of offspring are mat-ed, perhaps 10–15%. If all the offspring were mated, then we would lose the combination that had the highest fitness; there is no guarantee that any of the offspring will be better. The offspring that are chosen to be mated are randomly selected. Because there are more offspring for the fittest chromosomes, there is a greater chance that these will be used in the mating process.

Mutation

Mutation is the process of introducing new genes, or combinations of genes, into the gene pool. Up to this point, the genetic algorithm has propagated the fittest chromo-somes, then created new chromosomes by combining entire individuals of an existing one through mating. The last step is to randomly introduce new genes on the chance that an even better chromosome, or a better individual, will be formed.

A new gene can replace any one of the genes in the chromosome; however, it must be chosen from the corresponding list for that gene, as shown in the initial creation of the *gene pool*. For example, the first gene of the chromosome, which represents a trend-following calculation, can only be selected from the five choices designated. In mutating the gene, one of these five techniques is chosen randomly, excluding the exist-ing gene. Similarly, the calculation period and the way in which rules are combined are selected randomly from their corresponding lists. Typically, only one gene in one of the

individuals in a chromosome is mutated; however, you could mutate all of the genes in an individual. Using the approach of changing all the genes in one of the individuals in Chromosome 1,

Chromosome 1: MA, 10, <, C, [0], &, Stoch, 5, >, 50, 1

the genes in the second individual (Stoch, 5, >, 50, 1) might be mutated to (RSI, 10, >, 40, 1). Of all the chromosomes, only 10% are chosen at random for mutation. If too many are mutated, then there is a strong chance of losing the best chromosomes. If too few are mutated, then there is little chance of introducing a new and better chromosome. This same balance was discussed during mating.

The two methods of mating and mutation provide the only tools needed to introduce new features into the optimization and to combine features in all ways. If you could continue this process indefinitely, you would see every possible mutation and combination.

Converging on a Solution

Each pass through the genetic algorithm process creates a new generation of chromosomes. Normally, the population remains the same but the average fitness (the measurement of success) of the pool keeps increasing through propagation, mating, and mutation. The solution is found when either

1. The average fitness of the entire pool does not increase after a few new generations are created, or
2. There is no improvement over the chromosome with the best fitness.

Finding the best solution requires a fine balance of how many chromosomes are mated and how many of the new offspring are mutated. If only a small percentage (less than 5%) are changed, there is little chance of creating better offspring. In that case the results converge quickly because they are limited to choices and combinations within the initial pool. If a large percentage of the offspring (more than 25%) are changed, there is a significant possibility that the best of the pool will be lost and the process will fail to converge.

Putting It into Practice: Simulated Performance

The technique for a successful search using a genetic algorithm requires that good test procedures be followed. Data is separated into the testing and out-of-sample parts, where 70% to 80% of the data is tested. Once a solution is found, it is applied to the remaining 20% to 30% of the data not yet seen. If there is a large amount of data, the two sets could be 60%–40% or 50%–50%. The performance of the out-of-

sample data will normally deteriorate, but must still meet minimum success criteria. These expectations are discussed in the next chapter. Step-forward testing can also be applied. A complete discussion and analysis of testing procedures can be found in Chapter 21.

Multiple Seeding

Both the genetic algorithm and neural network are capable of finding a solution in much shorter time than a sequential search. In many cases, the solution can be found in seconds instead of hours; however, it may not be the best solution, just a very good one. Because there are so many combinations, you will never know whether the solution is the best—that would require testing every combination. It is most likely that the solution is a local peak rather than a global one. The analogy can easily be seen on a contour map. If you are dropped at random on one point of the map, and you proceed to walk uphill looking for a peak, you will find the local high point, but you may have missed a much higher mountain across the valley.

Because the propagation, mating, and mutation processes all seek higher fitness values, they essentially move uphill from the best combinations in the initial pool. It is possible that a combination of trends and indicators that would produce the best system was never introduced during the random initialization of the pool, nor did it appear in the subsequent mutation process.

The most efficient solution is to run the genetic algorithm or neural network a number of times to see if it arrives at the same solution starting with different random values in the pool. While it is not a guarantee that you will cover all major combinations, especially when there are a massive number of variations, it is the simplest way of avoiding serious oversights. The search tools are so fast that running them 10 times is still a minute amount of time compared to a sequential search. If you think about the contour map, starting at 10 different points has a much higher chance of finding the highest peak than beginning at one point.

REPLICATION OF HEDGE FUNDS

Technology allows us to approach problems as never before. With the proliferation of hedge funds, some analysts have decided that they can replicate the performance of those funds using mathematical methods, and any inaccuracy would be offset by the lower cost of replication compared to hedge fund fees. Replication can also be applied to a set of hedge funds, such as a fund of funds, or a hedge fund index. The systematic process of replication is marketed as a conservative and algorithmic approach to investing.

In a simple example, say that a hedge fund allocates resources only to stocks in the mining sector, and within that sector there are 10 stocks. The program is always long. If we have the historic returns of the hedge fund, we can use Excel's *Solver* and find what

percentage of the investment would need to be allocated to each of the 10 stocks in order to duplicate the returns. This assumes that allocations did not change over the period of analysis. It may be that positions are rebalanced once during the month, so that this regression may need to be done more often with some constraints on the size of the change allowed in each stock.

Alternatively, we try to replicate the returns each day by asking "What allocations would have given us yesterday's returns?" We then buy or sell shares to align with that solution. On the next day we ask the same question. If the hedge fund does not change positions often, we will adjust our positions until we align correctly with the hedge fund positions, similar to the way a Monte Carlo approach to sampling finds the right distribution. During the time it takes to find the same positions, there will be some tracking error, but this may be better or worse than the actual hedge fund.

The mathematics of replication has become a specialized area of application. Rebalancing of the replication model was initially done monthly; however, that would allow for large tracking errors. A better method would be to rebalance whenever the tracking error exceeded a threshold level. There is a trade-off between the tracking error and the cost of rebalancing. The Hedge Fund Replication Index (see *hedgefundreplication.com*) is said to be able to replicate 85% of hedge fund returns, leaving a tracking error of 15%, less than the fees associated with a fund of hedge funds.

System Testing

Numbers are like people; torture them enough and they will tell you anything.

—Anonymous (quoted by John Ehlers
in *Rocket Science for Traders*)

System testing remains a fight between computer power and common sense. Before 1980, testing a trading system was a simple concept. Conscientious traders and analysts applied their ideas to charts, marking the buy and sell signals manually and keeping a hand-written record of the trades. The process took a long time; therefore, they were careful not to begin unless they were reasonably sure that the method had a good chance of success. Ideas came from an understanding of the fundamentals or the awareness of market patterns, plus valuable experience. The tabulation of profitable results was verification that the idea was sound.

By the early 1990s, you were at the high end of technology if you had daily data that could be read by a computer and a program such as TradeStation or MetaStock to give you charts and indicators that could be combined into a trading system. Best of all, these programs showed the historic profitability of the trading strategies and, in particular, the best strategy. Moving averages and the RSI were still sophisticated tools, but the ability to test for the best calculation period and combine indicators began a step down a path that has now become very complex and taken the industry further from the simple solution.

That process has continued to advance, but not always for the better. Those who have applied diligence, selectivity, and most of all common sense to the automating systems are likely to have gotten better results, faster response, and more useful information than ever before. Those who have carelessly used the power of new computers and the convenience of trading strategy development software are not likely to be as successful. Bigger and faster computers and high-tech software do not mean better results.

The tools alone do not ensure a profitable trading program. It is the thought process, creativity, and knowledge that are needed to be competitive. Computers allow us to solve problems of magnitude and complexity that we could never have considered before, and they do it in a way that is easy to understand. We can manipulate massive amounts of data to find out not only whether our ideas are sound, but the risks associated with various events. But computers are still only highly sophisticated calculators, even when you call your technique "artificial intelligence." Giving it a catchy name does not make it true. Computers do not think, they compute.

One of the most misused techniques that has evolved from computerization is *optimization*. This is not the mathematical process of finding a *local maximum* on a map of continuous contours; instead it is the iterative process of repeatedly testing data to find the single best moving average speed, point-and-figure box size, stop-loss size, or a combination of multiple values that produce the highest trading profits, the lowest risk, or some other target result. That is not to say that optimization could not be effective. Traders seem to have an irresistible urge to know the system parameter values that would have generated the highest profits. Experience shows that the best test results never live up to their expectations and are often a financial tragedy when traded. They are the result of *fine tuning* or *overfitting*. Everyone using a computer for testing a strategy is guilty of overfitting the result at some time.

Computers present a serious, often futile, dilemma for the analyst. If the objective is to use a trend-following strategy—a moving average, exponential smoothing, or linear regression slope—how do you decide the right calculation period? More than 20 years ago, the 5-, 10-, and 20-day moving averages were most popular because they represented a 1-week, 2-week, and 1-month time interval. They were also easy to calculate. In fact, the 10-day moving average was the most common because no division was necessary to arrive at the average, you only needed to move the decimal point. Unfortunately, competition is now keener, markets more complex, and that simple approach is no longer viable.

Consider the process of creating a trading strategy. For a trend system, once the calculation period has been logically selected (e.g., 10 days) and the trading rules determined (buy and sell when the trend changes direction and use a 3% stop-loss), you will want to know the performance of that choice over past data for a selection of futures markets or stocks. If testing proves successful, you will trade using that method; however, if it does not meet expectations using the test data, what do you do? With a computer, the answer is easy: try another trend speed or a different stop-loss and see if the results are better. This follows a perfectly natural progression; after all, who would use a system that has failed to be profitable in the past?

Ultimately, we decide that testing all combinations of calculation periods and rules gives us a better understanding of how successful a trading system can be. Why guess at each parameter when you can try them all using specialized development software? We should know by now that this approach fails, but not everyone knows the approaches that are correct. The following sections discuss various aspects of testing and optimization, how to take advantage of it, and how to control it. There is a comprehensive summary at the end of the chapter.

EXPECTATIONS

Expectations are an essential part of system development. They force you to define a plan in advance. The importance of this comes when you see the test results.

Testing should be the process of validating your ideas. That first requires that you define those ideas in a clear plan, deciding whether it should work for hourly, daily, or weekly intervals. You should have an idea of the relative risk of the method, and the percentage of profitable trades. If the test results confirm your ideas, then you can have confidence in the strategy. If they differ, you know that something is wrong. Without expectations, there can be no validation or confirmation; therefore, if you indiscriminately test all combinations of indicators, you will not know whether you have found a good idea or have simply overfit the data. Testing should be used for deciding between various trade-offs of risk and reward, but *not for discovery*.

Setting Your Objective and Measuring Success

There are a number of steps that must be carefully performed before testing can begin. These steps are more important than the actual testing and will require you to decide what you are looking for and how you will use the results. Correctly defined, the final system will have realistic goals and predictive qualities. Incorrectly done, it will look successful but fail in actual trading.

First, define the *test objective*. What measurement or statistic will show success? Is it the highest possible profit, the frequency of profitable trades, or the average profit to average loss ratio? Testing software gives you a choice, but does not allow a combination of these items; most likely, it will be a combination that you want. For example, if maximum profits are used as the performance criterion, the resulting system may have one or two large profits and an overwhelming number of small losses. This can be seen in energy strategies that captured the move to $150/bbl in June 2008, or in gold where the 20 years from 1985 to 2005 were losses, but those were all forgotten when prices neared $2000/oz. When seen as total performance, exceptional profits during the peak of the stock market in the late 1990s can overwhelm the results of the whole test, making it look acceptable; however, the profits were made in the past and not the current market. That strategy may not be profitable now. Later, this chapter will look at successful but decaying performance.

While there is no *right* way to define success, the industry has favored the *Sharpe ratio*, or the simple *information ratio* (without the risk-free returns)

$$Sharpe\ ratio = \frac{Annualized\ returns - Risk\text{-}Free\ returns}{Annualized\ risk}$$

This and other general performance measurements can be found in Chapter 23. Later in this chapter, the section "Performance Criteria" explains many of the components that can be used to decide the relative success of a trading method. All traders want to know

the maximum drawdown, not just the probabilities, and day traders will want to know the average profit per share or profit per contract to understand the impact of costs and slippage on the results.

IDENTIFYING THE PARAMETERS

Once the test strategy is known, the *parameters* to be tested must be identified. A *parameter* is a value within the strategy that can be changed in order to vary the speed or timing of the system. For example, parameters include:

- Moving average calculation period (number of days or bars).
- Exponential smoothing constant (a percentage).
- Band width around a moving average (percentage or number of standard deviations).
- RSI calculation period (number of days or bars).
- Stop-loss value (fixed number of points, a percentage, or a volatility factor).
- Maximum holding period for a trade (days or bars).

Suppose the strategy uses a moving average to identify the trend, an RSI to decide the entry timing, and a percentage stop-loss. The most basic test would be to choose the stock or futures contract, fix the calculation period of the RSI at 14 days, fix the stop-loss at 3%, and vary the calculation period for the moving average. The test would then simulate the results of the trading strategy, stepping through the possible range of moving average days: for example, 1 through 250 days. This is called a *1-parameter* optimization.

Most systems have more than one important parameter. In this example, the moving average period, the RSI period, and the stop-loss value are all important. If the stop-loss is held constant at 3%, the moving average and RSI can be tested together. The optimization process does a sequential test, selecting the first moving average value, then testing all of the RSI values; the next moving average value is then used, and all of the RSI values are retested. The process is repeated until all combinations of the moving average calculation periods and RSI calculation periods have been tested. This is a *2-parameter* optimization.

Range of Parameter Settings

It doesn't take much effort to create an optimization with a very large number of tests; it follows that the test time for the optimization will increase dramatically. For example, a test of 250 moving average values (1 to 250), 30 RSI calculation periods (5 to 34), and 30 stop-loss values (0 to 3.0 in steps of 0.1) gives $250 \times 30 \times 30 = 225,000$ tests. If each test takes 0.01 second (much longer for intraday data, perhaps less for daily data), the total test time will be about 36 minutes. Tests with more parameters will quickly get out of control. It will be necessary to select fewer values to test for each parameter and

choose them more carefully. Limiting the test is actually an advantage because it forces a decision, in advance, about the range of parameters that are reasonable for the strategy. Without these restrictions, the test process moves further away from validation and closer to indiscriminant exploration.

Distribution of Values to Be Tested

It is not necessary to test all of the calculation periods for a moving average or any trend method; however, the values should not be spaced evenly apart. For a moving average trend, it is best to use all of the smaller values for the number of days and then fewer values as the numbers become larger. For example, the sequence

$$1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 20, 25, 30, 40, 50$$

would be a better choice than using all numbers between 1 and 50. Consider that the difference between a 49- and 50-day period is only 2%, while the difference between the 2- and 3-period test is 50%. The tests at the two ends of the test series therefore represent very different sensitivity. The difference in performance between a 1- and 2-day moving average will be significant, but the difference between a 49- and 50-day trend will be negligible.

Using equal increments will weight this set of tests heavily toward the long end; that is, average results or a plot of test returns versus calculation period will always reflect the results of the longer calculation periods. Using test values that increase in percentage increments would solve this problem. If the starting test value is multiplied by 1.5 to get the second value, and that is multiplied by 1.5 to get the third value, the final test series is improved. Beginning at 1, the series becomes

$$1, 1.5, 2.2, 3.3, 5.0, 7.6, 11.4, 17.1, 25.6, 38.4, 57.7$$

By specifying the tests using percentage increments you not only reduce the test time but you improve the distribution of results.

Ordering the Test Parameters

Testing one parameter at a time can give you a better understanding of the dynamics of a system. It is much more difficult, sometimes impossible, to see the way that test results vary when there are five parameters being tested simultaneously. For sequential testing, the most important feature must be tested first—the one that combines the big picture with the greatest impact on the results. In the example of a system with a moving average, RSI, and stop-loss, the moving average is the primary component. The success of the strategy will depend on selecting a profitable trend period. While the other indicator and risk control should improve the results, it is unrealistic to think that a strategy should be driven by the size of the stop-loss. In the same way, the second test should be the RSI to improve entry and/or exit timing. Risk control, using the stop-loss, should improve the final picture.

Types of Test Variables

Test variables can be of three forms: *continuous*, *discrete*, or *coded* (alphabetic). In order to interpret and display test results properly, these forms need to be identified in advance.

> *Continuous parameters* refer to values, such as percentages, which can take on any fractional number within a well-defined range. If a stop-loss level is defined as a percentage, it may be tested beginning with the fractional value .02% and increasing in steps of .005% until 2.0% is reached.
>
> *Discrete parameters* are whole numbers, or integer values, such as the number of days in a moving average.
>
> *Coded parameters* represent a category of operations, also called a *regime*. For example, when the parameter value is *A*, a single moving average is used; when the value is *B*, a double moving average system is tested; and when the value is *C*, a linear regression is used.

The first two parameter types, continuous and discrete, should be distinguished from the last one, coded. The analyst can expect a pattern in the results when testing parameters that take on progressively larger or smaller values. For example, a moving average of 1 day may post the largest losses, and show progressively smaller losses until 30 days is reached, where it becomes profitable; performance peaks at 80 days and then starts to decay. The coded parameter, however, shifts rules for each value. There is no reason why a change of rules, which switches regimes, would result in any performance pattern that continues across these regimes. The first rule may be profitable, the second losing, and the third profitable. The display of results, discussed later, is only valid for continuous and discrete parameters tested in an incrementally ascending or descending manner.

SELECTING THE TEST DATA

It is assumed that strategy testing will be done using a computer. It will be necessary to have a database of prices and if more ambitious, a history of economic statistics as well. Data is readily available from numerous vendors and includes daily high, low, and closing prices, plus volume and open interest (futures) for nearly all exchange-traded markets in the world. Individual stock data is also available and may be found on the Internet at no cost (although some inconvenience). Intraday (tick) data on futures can be purchased from a limited number of vendors, the most well-known being CQG Data Factory and Tick Data, Inc. Stock data can usually be exported from whatever quote service you use. A complete database is an important asset. It can be kept current by automatic updating at the end of each day.

Some data vendors, aware that analysts will want to use the data for testing, have created special *continuous test series* specifically for that purpose. Prices of individual
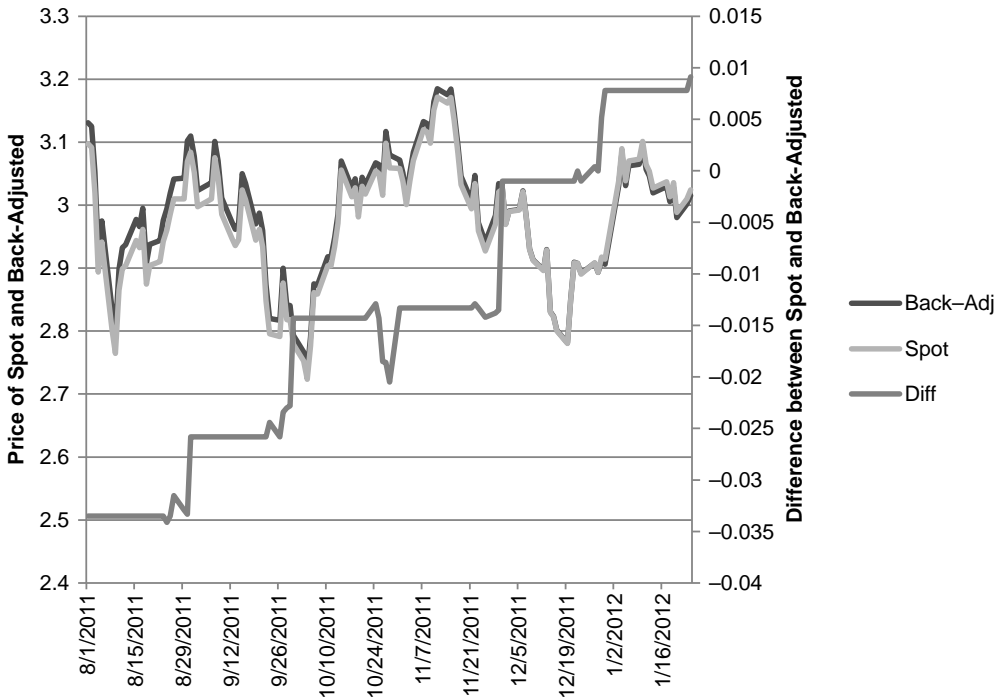
FIGURE 21.1   Comparing back-adjusted data with the nearest contract.

stocks will be corrected for stock splits by dividing the data prior to the split by the split ratio; therefore, a 3:1 split will be adjusted by dividing the data before the split by 3. If there was a previous split of 2:1 the data prior to that is divided by 6.

Futures data, which expires every 1 to 3 months, is rolled by adjusting the gap between contracts on the date of the roll. Vendors allow you to specify exactly how and when the rolls should occur, or you can simply select the default method. Figure 21.1 gives an example of a recent period of heating oil futures, which rolls every 30 days. Over a short time period, the unadjusted and adjusted series look very similar, but the differences (Diff) show that the gaps cause sharp price changes each month, and the accumulated gaps over 10 or 20 years of data can be enormous. Without the gap adjustments, these differences become windfall profits and losses. Analysts may choose to use different data formats for testing.

1. *Individual full contracts.* The entire futures contract or a series of contracts is tested. Using full contracts is a tedious process and can result in overlapping test periods and duplication of results. Figure 21.2 shows the test bias resulting from carelessly using all of the data for futures contracts from July 1993 through December 1996. The horizontal bars represent the life of the contract for a futures market with delivery in May, July, September, December, and January. Because of the 18-month duration, there was only 1 contract tested in the last half of 1991, compared to 6 contracts
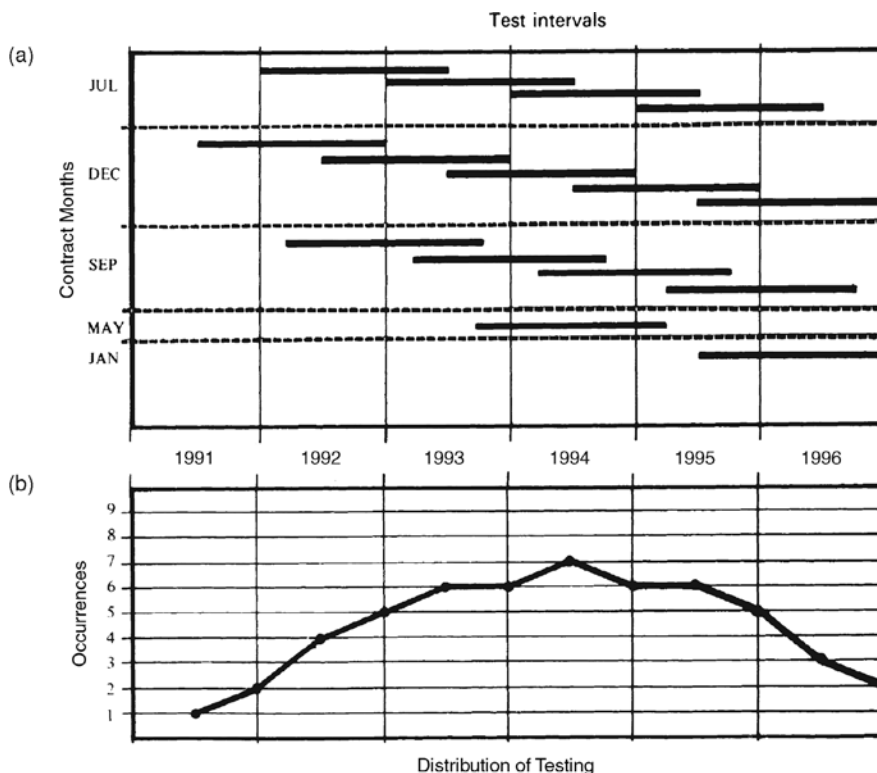
FIGURE 21.2 Testing using all of the data for futures contracts July 1993 through December 1996. (a) Test intervals. (b) Distribution of testing.

tested during the middle period, and falling off to 2 contracts at the end of 1996. Figure 21.2b shows the frequency of testing for each time period.

2. *Backward- (or forward-) adjusted data.* Individual contract data segments are combined on a specific *roll date* by adjusting all the prices in the older contract by the difference between the prices of the new and old contract on the roll date. For example, if on the same date the new contract was trading at 80.00 and the old contract at 75.00, then all the data in the old contract would be increased by 5.00. Moving backwards to the previous contract, if that showed a price of 78.00 when the third oldest contract showed 74.00, the prices in the third oldest contract would be increased by 9.00, the sum of the two roll values. Had the third oldest contract been higher than the second on the roll date, 81.00 versus 78.00, then its prices would be increased by only 2.00, the sum of 5.00 minus 3.00. The adjustment in every older contract is the cumulative adjustment of all previous roll values. A *forward-adjusted* series begins with the oldest data and adjusts newer data as it moves forward in time. At the end, the old data matches the actual prices, but the newest data is different.

This technique has become popular and works well for most systems. The adjustments will cause large differences in the oldest data, especially if that data goes back more than 10 years. In the extreme case, the adjusted price can even become negative! Calculations that use a percentage of price do not work with back-adjusted data.

The *roll date* can be one of the following choices:

a. *A fixed date*, such as the last day of the month before the delivery month. This is standard for interest rates, agricultural products, and some other commodities because margin requirements change significantly for those who trade into the delivery month. Therefore, for Treasury notes and Eurodollars, the June contract is rolled on the last day of May, the September contract on the last day of August, and so on. For index markets, the expiration date may be the third Friday of the month, and the contract is actively traded until a few days before delivery. Software provided by data vendors allow you to choose n days before the contract expiration.

b. The day on which *volume and/or open interest switches* from the nearest delivery to the next deferred. Using this method, the roll-adjusted series represents the contracts with the greatest liquidity.

3. *Indexed data series.* A continuous series is constructed to represent the daily percentage returns each day. This is essentially the way a stock index is constructed. If the price of copper moves 1.0% from $0.9000 to $0.9090 per pound, then the indexed value will be increased by 1.0%, from 100 to 101. When the contract rolls, the price change of the new contract is compared to its prior price. The conversion into an indexed series is useful for comparing two markets and for constructing portfolios.

4. *Constant n-day forward series.* In the same manner as the London metals and cash forward markets are quoted (for example, delivery 30 days forward), an artificial 30-day forward series can be created for any futures market. The 30-day price is calculated by interpolating between the delivery months on both sides of the date 30 days ahead. This type of data series is called a *perpetual contract* and presents its own set of problems. Deferred contracts are not as volatile as the nearby contract, and formulated values create a series that tends to be smoother than the one you are likely to trade. Your test results may appear to have less risk, probably less profit, and greater reliability in identifying trends. Most important, testing would be done on a price series where none of the prices actually occurred.

Cash Data

Cash data is available for most futures markets, but FX presents a particular problem. There are many more FX crossrates than there are futures markets, and some of these can be interesting to trade. Even the Euroyen, the second or third most active trading FX market, has very low futures volume on the CME; therefore, the data may not be as

good as cash data. However, testing cash data and trading futures data is not a reliable approach. Futures prices can trade at a premium or discount to cash and will fluctuate with the difference in interest rates between the two countries as well as speculative anticipation. In the extreme, the cash market may be going up and the futures market going down. Approximate futures data can be created from cash data by calculating the interest rate parity, the forward difference in the interest rates of the two component countries in the crossrate. See Chapter 2 for more detail.

Problems with Adjusted Data Series

Back-adjusted data, split-adjusted data, and indexed data are ideal for testing most methods, in particular trends and momentum indicators. They cannot be used if any of the calculations require percentages. Futures markets, especially energy, have large price differentials on the roll dates. This causes older prices to be shifted down so that the prices at the beginning of the adjusted series can be negative. This makes no difference to a moving average calculation, but a percentage of a negative number is likely to play havoc with the strategy's trading decisions. This problem is more subtle for back-adjusted stock splits because the prices all appear to be scaled correctly and never go negative. For example, if IBM had split twice (each 2:1) during the past 10 years, then the price 10 years ago would have been adjusted lower by dividing by 4. If the real price in 1990 was $50 per share, the adjusted series would show it as $12.50. If the trading range was $2 at $50 per share it will be $0.50 when back-adjusted to $12.50. But the volatility at a truly lower price might be proportionally higher than at $50. The $12.50 stock will have all the characteristics of a $50 stock.

Comparing historic highs and lows cannot be done with a back-adjusted series. Analysts may normally recognize an important resistance level when prices test the highs of the previous year; however, the older price has been adjusted and may be higher or lower than its original price. To make this comparison, it is necessary to have a second series that is rolled on the same dates but not back-adjusted. The second series can be accessed when the true historic price is needed.

An indexed series can be awkward when you trade because all the prices must be converted to indexed values in order to be evaluated by the strategy. You cannot simply test using an index series, then trade using a price series.

Which Data Series Is Best?

If all adjusted series have problems, which is best? It is the series constructed from the pieces of the contract that you actually trade. In most cases, it is a back-adjusted series rolled on the volume or open interest. It is also the most liquid. If the true prices are needed, then a second series should be used.

Indexed series are also very good, although more difficult to trade. The process of indexing comes very close to volatility-adjusting and may make the data easier to use and the parameter selection more consistent.

Synthetic Data

The ultimate answer to getting test data is to create it yourself. The more data the better. It would be ideal if *synthetic data* could be created that had the same characteristics as the data that will be traded, yet is different. Over the years analysts and mathematicians have attempted to create synthetic data.

> ***Monte Carlo Method***. One of the first methods used a Monte Carlo process that first scanned actual data looking for a way to divide the data into equal-length segments by finding the period that resulted in the lowest autocorrelation (the dependence of one data segment on the previous). The real data could then be rearranged randomly into a new series. Unfortunately, mixing segments would separate a prolonged bull market from the inevitable decline that occurs at the peak. In actual trading, a system can suffer a large loss when prices peak only after it has accumulated larger profits during the previous upwards trend. Separating the data so that these two events do not follow one another is unrealistic. The Monte Carlo approach was abandoned.
>
> ***Random Numbers and Distributions***. A synthetic series created mathematically begins with random numbers. It has been shown that a simple series of random numbers will look very similar to most price movement, especially during a sideways period. But price movement is not normally distributed and has much longer runs, both up and down, than would be found in a random series. Real data may have seasonality, a long, underlying directional bias due to government policy, changing volatility, and a certain level of underlying noise. But mathematics is powerful, and an artificial series can be created that has 10% trend and 90% noise, reflects a seasonal bias, and has a lognormal distribution.

A more recent and sophisticated attempt at creating synthetic data can be found in Wolberg's *Expert Trading Systems*.[1] The technique uses a Gaussian distribution to represent the skewed price distribution, and a noise component that is a function of the price trend, yet variable around the trend. The trend component is itself nonlinear, created using kernel regression. One process used for verifying that the synthetic data has the same characteristics as real data is the *Fraction Same Sign*, which is the percentage of synthetic records in which the predicted and actual data have the same sign. Other statistics of variance, noise, and distribution should match.

Synthetic data is an ideal solution to a difficult problem. It also brings to mind two computers playing chess with one another. In this case, one computer creates a data series to fool the other computer into thinking that it is real. The second computer is creating a strategy that can take advantage of the fact that it knows the data series is not real, and is constructed from components that are built systematically. If the series was

---

[1]John R. Wolberg, *Expert Trading Systems: Modeling Financial Markets Using Kernel Regression* (New York: John Wiley & Sons, 2000).

real, then it would have a level of complexity that resists exploitation. If it is not random, then another computer should be able to identify the process used to create the series, then capitalize on it. The second computer should always win.

Traditionally, analysts use as much data as possible to determine whether their proposed strategy is sound. The use of long test periods assures a good sample of markets, including long periods of sideways movement, bull and bear markets, and a good number of price shocks of various sizes. As more data is used, it is most likely that the results will show larger profits and larger losses. More data allows the possibility of larger price runs, larger price shocks, and more complex patterns.

At the risk of offending someone, the testing process most often involves some level of cheating. Acknowledging that developing a system can be difficult, there is a right way and a wrong way to test in order to arrive at credible results. This section will outline the most acceptable methods.

In-Sample and Out-of-Sample Data

The basic scientific method begins with *in-sample* data for testing the idea. In-sample data is a part of the total data available that has been set aside for testing. Normally, it is about 50% of the data. The remaining part of the data will be used for *out-of-sample* testing, or validation. If, for example, there are 20 years of data, the in-sample data could be

1. The first 10 years.
2. The last 10 years.
3. Alternating, equal periods, such as the first 5 years, then years 11–15.
4. Alternating random time periods.

Setting aside the first 50% of the data is very common, but it suffers if the markets change, becoming more or less volatile or changing characteristics due to market dynamics. The same is true of using the last 50% of the data. Choosing parameter values that worked in older or newer markets is not a robust approach.

Alternating periods gives a better sample of changing markets, and good results will require some type of adapting rules. When using alternating periods, the system actually calculates performance on all data but only displays the in-sample periods. Results should always be shown as daily changes, not total profits. If you choose total profits, then the difference between the end of one in-sample period and the start of the next will show whether the trading method netted a profit or loss during that interval. That would be cheating. Instead, the daily returns can be combined into one continuous stream and analyzed.

Alternating random periods are the most robust testing method, but those periods must be fixed at the beginning of the development process and not changed. If you were working on two different systems, each one would have its unique set of test data. In that way, you are not always testing the same data for every method you create. To create random test periods requires a small program, but that could be implemented on a spreadsheet.

- Decide how many in-sample periods and multiply by 2. Let's say there are only 2 in-sample and 2 out-of-sample periods
- Does the in-sample come first? Let's assume *yes*.
- Then, for 20 years of data, each segment will have about 5 years. Let's say that the shortest will be 2 years and the longest 8 years.
- Start the first in-sample period on day 1.
- Find a uniform random number (between 0 and 1) and multiply that value by 6, the range of years in the in-sample period. If the random number is 0.72 then the adjusted value 4.32 is added to the 2-year minimum giving the length of the first period as 6.32 years.
- The same calculation is repeated for the next two segments. The last segment is whatever is left over.

Once calculated, these intervals can be assigned to specific data and all in-sample testing will use that data. This data can be used as much as needed. Rules and parameters can be changed, and a detailed inspection of performance should be done. The data can be tortured without mercy. Evaluating these results will be discussed in the next section.

Out-of-Sample Testing

When testing is completed, it is time to validate the results using the out-of-sample data. This is the critical part. You must have done all you can to create a good system before using the out-of-sample data, because there is no second chance.

It is best to use only the out-of-sample data in your final test, and not include the in-sample data. Rarely, it will have the same successful profile as the in-sample test. More often, it will show about 50% of the returns from the in-sample data. Because there is 50% new data containing price moves never seen in the original testing, the performance is often worse. A drop in the information ratio from 2.0 to 1.0 for the out-of-sample data is actually a reasonably good result. But what happens when the ratio drops from 2.0 to 0.4? When the detail is inspected, you realize that you overlooked a case where the volatility was extreme, and adding a simple volatility filter would fix it, increasing the ratio to 1.0. This is where *integrity* enters. You cannot fix anything once you have used the out-of-sample data. That is called *feedback*. The result is always overfitting, and the chance of the system working plummets. You can spend endless time using the in-sample data, but once the out-of-sample is used, you're done.

Step-Forward Testing

The repeated process of choosing parameter values from tests of in-sample data, then applying the results to out-of-sample data, has been developed into a process called *step-forward testing*[2] (also called *walk-forward testing* and *blind simulation*). In many ways this parallels what happens in reality. We choose the parameters on back data, then use them going forward. This process goes as follows:

1. Select the total test period—for example, 20 years of daily data from 1988 through 2007.
2. Select the size of the individual test intervals—for example, 2 years.
3. Begin testing with the first 2-year period, 1988 through 1989.
4. Select the best parameter values from those results.
5. Use the selected parameter values from Step 4 to find the performance of the next 6 months of data (the first half of 1990). Accumulate this out-of-sample performance throughout the test process.
6. If there is more data to test, move the test period forward by 6 months (the second test will be from the second half of 1988 through the first half of 1990), then test the data; otherwise, go to Step 8.
7. Repeat Step 4.
8. The result of the step-forward test is the accumulated results of the individual out-of-sample tests in Step 5.

This process clearly simulates the traditional approach to test design, using in-sample and out-of-sample data in its proper order. It seems intuitively satisfying.

Short-Term Bias

Step-forward testing can introduce a bias that favors faster trading models. Because the in-sample test used only 2 years of data in our example, a long-term trading model may only post a few trades during this period. If there was a bull market during the entire 2 years, then any long-term trend would have captured the same results. Without enough data, testing cannot find which calculation period is best. Consequently, it would be difficult to evaluate the risk of these positions.

The problem of short-term bias can be identified by the optimum parameter values switching from short-term to long-term in successive test periods. For example, if you test moving averages from 5 to 50 days, you will find that one period shows 10 days as the best, followed by a period in which 50 days is best, followed again by a period in which 15 days is best. This erratic behavior is a sign that the data intervals are too short. It can

---

[2]A thorough presentation of both step-forward testing and optimum search techniques can be found in Robert Pardo, *Design, Testing and Optimization of Trading Systems* (New York: John Wiley & Sons, 1992).

be corrected by increasing the test intervals from 2 years to 5 years. Longer test intervals go a long way toward eliminating the short-term bias, but they must be considered in context with the type of trend envisioned by the developer. It also requires extra data at the beginning that cannot be used in the final results.

Feedback

The most serious concern about step-forward testing is feedback, discussed in the previous section on out-of-sample testing. Each application of the step-forward process uses all of the data, where the final performance is determined from the out-of-sample part. Then, if the test is performed more than once, there is no out-of-sample data, which violates our concept of scientific testing. To avoid this, additional data can be set aside for a final test, but that reduces the amount of data available for testing and validation.

SEARCHING FOR THE BEST RESULT

The best result, or *optimal result*, of testing a strategy is the *most robust* set of parameters. That is different from the parameter set that yields the *maximum value* of the test objective, whether it is highest profits, most consistency, or highest information ratio. It unlikely that the maximum value will be the selection that represents the most robust strategy; however, many methods require that maximum be located in order to decide the most robust. This section will discuss some of the techniques used to find the maximum result.

Search methods have evolved with technology over the past 20 years. At the high end are genetic algorithms and Monte Carlo sampling that are used when test combinations are too large to scan all possibilities. They will be discussed toward the end of this chapter. The most basic approach, *sequential testing*, calculates the results of every combination of parameters in a preset order. For example, a single moving average system with a stop-loss might have the following parameter test ranges:

1. Moving average calculation periods from 1 to 50.
2. Stop-loss values from 0 to $1,000 per share, or the equivalent, in steps of $10.

The total number of tests would then be $50 \times 100 = 5,000$. When three parameters are tested, the number of tests increases rapidly. Taking care to use the best parameter distribution, as discussed earlier, is a simple way to reduce the test time and improve the usefulness of the results. When there are too many tests, even with carefully chosen values, a wider spacing of values can be used to locate the general pattern of performance. Once the approximate range of parameter values is known, more detailed retesting can be done.

Calculate the number of tests you will run before beginning an optimization. Although the time may not be a problem as computers have gotten faster, you may be limited by

available storage space. All computers have limits. If you generate 10,000 combinations, remember that you will need to evaluate them using a spreadsheet. Large numbers of rows and columns, combined with numerous markets, can create an unmanageably large spreadsheet.

Mathematical Optimization

If you have created a system on a grand scale and the number of parameters combined with the number of tests for each parameter generates a test that is far too long, then *mathematical optimization* offers a testing solution. One caveat is that this method only applies to continuous results. Test values must be increased incrementally, as in a moving average calculation period, with results that are also continuous; that is, performance increases and decreases somewhat smoothly as parameter values are increased. You cannot use this method to test coded variables that cause an abrupt shift in the methodology. In the simplest example, the procedure is:

1. Test one parameter, such as the moving average calculation period. Initialize the other parameters to a reasonable value.
2. Find the best result and fix that value for the remaining tests.
3. Select the next parameter to test.
4. Continue at Step 2.

This technique is called *sequential optimization.* It has the advantage of reducing a 3-parameter test case from $n_1 \times n_2 \times n_3$ total tests to $n_1 + n_2 + n_3$. It has the disadvantage, however, of not always finding the best combination of parameter values when the test is very large. Figure 21.3 shows a 2-dimensional map of test results, a good way of visualizing the relationships between parameter sets. If Figure 21.3 shows all the tests, then the best choice is the 20-day moving average with a 0.30% stop-loss (shown as 30 basis points), netting a score of 400. Using sequential testing, beginning with the moving average days, and setting the stop-loss to its lowest value, the best result (in row 1, at the top) is 200, the 15-day calculation period. Fixing the first parameter value at 15 and testing the stop-loss values, the results appear in the same column. The best result is a score of 300 when the stop-loss value was 0.30%. While this result is good, it is not the peak result or even next best, 350. For this set of tests, the best results cannot be found regardless of whether the test begins with rows or columns—when the lowest value for the other parameter was chosen.

There are two ways that avoid this problem. One method is to list the parameters in order of importance, that is, those that would have the greatest impact on profitability, or by its fundamental importance to the strategy. This was discussed in the section "Ordering the Test Parameters." In this example, the moving average is the primary parameter; however, that choice does not work. Fortunately, mathematicians are prepared for this problem.

Moving average days

Stop-loss (points)	5	10	15	20	25
10	100	150	200	150	150
20	200	150	250	300	300
30	250	200	300	400	350
40	200	150	200	300	300
50	100	100	150	250	250

FIGURE 21.3   Visualizing the results of a 2-dimensional optimization.

The other alternative is to choose a series of random starting values for each of the parameters, including the primary one, and search horizontally and vertically from that point. For example, two random values, 1 and 4, are chosen for parameter 2, the stop-loss. The test begins by scanning row 1 for all values of the moving average, which was done in the first example. The maximum score was 300. The test is repeated, scanning row 4 for all values of the moving average. This time the value 300 is found in column 4. The next scan of column 4 finds the maximum, 400. Instead of testing all combinations, $n_1 \times n_2$, the answer was found searching $2 \times (n_1 + n_2)$. This is much faster when n_1 and n_2 are large. This process is called *seeding*.

The object of any search is to find the best results with as few tests as possible. Although each 3-parameter test is only $n_1 + n_2 + n_3$ tests, the total number of tests will accumulate quickly. When you have performed a number of searches beginning with a random seed and found the same results, you can conclude with some certainty that you have located the peak. For example, if you begin searching in Figure 21.3 in any row other than row 1, or in any column at all, you will find the maximum value. For a test of this size, after 3 seeded values show the same result, you can conclude that the peak has been found. For larger tests, more seed values would be needed, yet the total number of searches will be relatively small.

More on Test Continuity

Any system can be tested using a computer, but automatic selection of the best results only makes sense when they are continuous. A moving average system with stops is continuous in all directions. Adding a third dimension, an entry or exit delay (0, 1, or 2 days) will produce continuous results in three dimensions. But not all tests satisfy this

principle. It has been briefly mentioned that rule changes are the most likely cause of discontinuity. Consider a test where one variable is the trend calculation period and the other is a set of rules, coded as *L*, *S*, and *B*, where *L* only takes long positions, *S* only takes short positions, and *B* takes both long and short positions. Although there may be interesting patterns and relationships, averaging these results does not make sense.

If each parameter tested has values that result in a continuous return curve, there are mathematical techniques available to locate the point of maximum profit with the minimum steps. One method, *optimization by steepest descent*,[3] uses both the returns and the rate of change of returns to determine the size of the steps taken during the optimization process. This method can be suboptimal if there is more than one return peak, but can be solved using seeding.

VISUALIZING AND INTERPRETING TEST RESULTS

Visualization is vital for avoiding gross errors in development. Looking at a price chart of what is being tested avoids data errors, and presenting test results in some logical form shows outliers, erratic results, or various levels of continuity. Some will argue that this can all be done by scanning the results and using various measurements, yet there is no real substitution for what the eye can see.

Test platforms, such as TradeStation, MetaStock, and Ninja Trading[4] are excellent alternatives for analysts who do not want to develop their own test platforms. TradeStation performs an optimization by sequentially processing every combination of parameters from a range specified by the user. The results are shown as a table, which can be exported to Excel, with one combination of parameters following another and the results of the test displayed on one line. As an example, if we are testing a moving average with a percentage band, an optimization was specified with

1. The moving average calculation period varying from 10 to 50 days in increments of 10 days.
2. The percentage band (based on the moving average value) varying from 1% to 4% in increments of 1%.

The system buys when the closing price penetrates the upper band and sells when it penetrates the lower band. The results of the optimization using NASDAQ 100 continuous futures from 1996 through 2003 are shown in Table 21.1. The user can choose the statistics that appear in the table. In this test, the columns (from left to right) are the moving

---

[3] F. R. Ruckdeschel, *BASIC Scientific Subroutines*, vol. 2 (Peterborough, NH: Byte/McGraw-Hill, 1983).

[4] There are an increasing number of test platforms available for a wide range of costs. Many of these will advertise in, or have been reviewed by, *Futures* magazine, *Technical Analysis of Stocks & Commodities*, *Active Trader*, and *The Technical Analyst*.

TABLE 21.1 Optimization Report from TradeStation for a Moving Average—% Band System Applied to NASDAQ Continuous Futures

MA Per	% Band	Net Profit	Long NetPrf	Short NetPrf	#Trds	#PftTrds	Avg W/L	Avg Trade	Avg Wbars	Avg Lbars	MaxDD	PFact	ROA
10	3	−89895	−32540	−57355	89	30	2.03	−1010	42	12	−189510	0.884	−47.4
10	4	−243235	−127165	−116070	61	36	1.23	−3987	46	16	−421200	0.693	−57.7
10	2	−194545	−89410	−105135	136	32	1.66	−1430	25	8	−227740	0.792	−85.4
10	1	−108775	−48005	−60770	192	35	1.56	−567	18	5	−157850	0.876	−68.9
20	2	−158015	−72245	−85770	94	32	1.61	−1681	39	10	−338560	0.792	−46.7
20	4	93265	52115	41150	48	41	1.69	1943	62	21	−120550	1.209	77.4
20	3	118545	67580	50965	60	41	1.73	1976	53	15	−221180	1.238	53.6
20	1	−39195	−13215	−25980	134	31	2.07	−293	29	7	−153890	0.945	−25.5
30	3	89365	51445	37920	52	32	2.48	1719	69	19	−128395	1.202	69.6
30	4	−12345	−13990	1645	43	32	2.02	−287	73	23	−210745	0.974	−5.9
30	1	178315	95540	82775	96	33	2.7	1857	43	8	−164980	1.352	108.1
30	2	65770	39605	26165	68	32	2.37	967	55	15	−141660	1.131	46.4
40	3	198085	103005	95080	43	34	2.9	4607	79	21	−150670	1.555	131.5
40	2	285980	152810	133170	51	35	3.46	5607	74	16	−154510	1.885	185.1
40	1	200735	106750	93985	78	28	3.79	2574	58	11	−147845	1.489	135.8
40	4	176195	80480	95715	31	38	2.37	5684	91	31	−161050	1.496	109.4
50	2	197125	107520	89605	49	34	2.89	4023	78	17	−116830	1.537	168.7
50	1	267520	140480	127040	56	33	3.51	4777	75	12	−90680	1.805	295
50	4	225465	106965	118500	27	40	2.57	8351	111	29	−152850	1.764	147.5
50	3	247025	116200	130825	31	38	2.95	7969	98	27	−127050	1.865	194.4

average period (*MA Per*), the percent band (% *Band*), net profit(*Net Profit*), net profit for long positions (*Long NetPrf*), net profit for short positions(*Short NetPrf*), number of trades (*#Trds*), number of profitable trades (*#PrfTrds*), average win/loss ratio(*Avg W/L*), average profit per trade(*Avg Trade*), average number of bars in a winning trade (*Avg WBars*), average number of bars in a losing trade (*Avg LBars*), maximum drawdown (*MaxDD*), profit factor(*PFact*), and return on account (*ROA*). *Profit factor* is the ratio of gross profits to gross losses, and *return on account* is the ratio of net profits to the absolute value of the maximum drawdown.

The optimization report is shown in the order that it appears after testing is completed. The moving average period increases from 10 to 50; however, the % band is not always in ascending order. Nevertheless, there is a clear pattern that shows profits begin with a 30-day moving average using a 2% and 3% band, then become profitable for longer calculation periods with all band values. When the optimization is completed, a final report is available showing the returns for the best combination of parameters, the 40-day moving average with a 2% band.

Seeing Continuity in Test Results

Visualizing the test results properly can make interpretation far easier. Table 21.1 is easy to read for the short test example, but combinations of 2 or more parameters can produce hundreds of lines of statistics that are not easy to organize. Figure 21.3 has already given a preview of how results might be presented in 2 dimensions to show continuity, rather than as a list of one test result after another, one line under the other. Consider the most basic test, a 1-parameter optimization of the moving average calculation period. Figure 21.4 shows three possible patterns that might result.

In all three patterns, the total profit/loss is shown on the left, and the moving average calculation period is shown along the bottom. Figure 21.4a indicates that both fast and slow parameters generate losses, although the midrange is profitable. It is easy to select the number of moving average days that produced the highest profit because of good continuity around that point.

Figure 21.4b shows two possible areas of success. Because both are about equal in size, the choice between the two would be based on frequency of trading and size of the profits per trade. The short calculation period has more trades, a shorter holding period, and smaller profits per trade. Another possibility is to trade the moving averages represented by the peaks. If you must choose one, then it is most likely that the longer calculation period will be considered a more conservative selection. A fast trading system can result in errors, poor execution prices, and large transaction costs. Unless these factors have been incorporated into the trading strategy, a highly profitable simulation may result in disappointing real performance.

The pattern in (b) appears exaggerated in (c), but is actually more realistic. The fast system has much higher profits in a very narrow range while the long-term trend has a wide range of success. The spike in the area of the 3-day moving average is surrounded by losses, indicating that the high profits may have been caused by a short-lived price
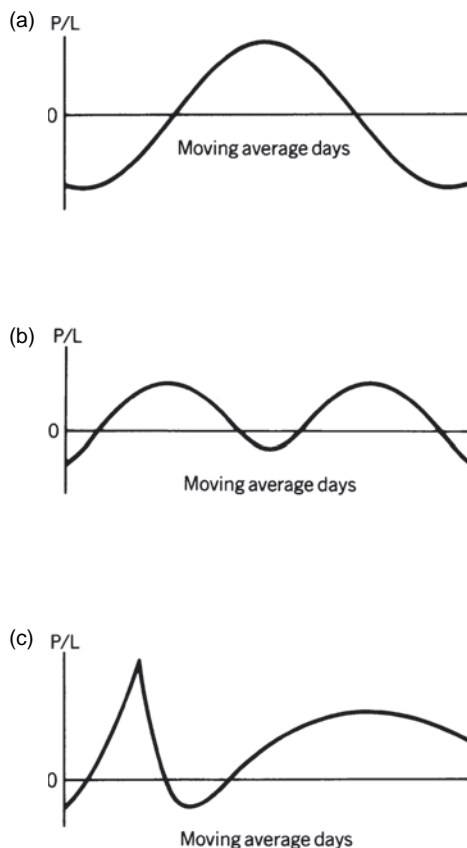
FIGURE 21.4 Possible resulting patterns from a test that varies only the moving average days.

shock or an actual window of opportunity for a fast trader. This pattern also occurs when there are no transaction costs. There are fast trading strategies that can be highly profitable without slippage or commission! The pattern in (c) is also exaggerated by testing calculation periods with equal increments. As discussed earlier, calculation periods of 90 and 100 days will have similar results compared to periods of 10 and 20 days.

The three patterns in Figure 21.4 are all good results, showing a wide range of calculation periods that produce profits. An erratic pattern, one that does not have these smooth return relationships, shows that the underlying strategy is unreliable, or that the test program has errors.

Two-Parameter Tests

The most common tests are those that have two or more parameters, a combination of one or two trends, an indicator for timing, a stop-loss for risk control, and sometimes profit targets. Table 21.1 showed the results of a moving average with a percent band.

The results of a 2-parameter test, for example, a moving average and a stop-loss, can be shown as a *grid* (a 2-dimensional table) in the same form as Figure 21.3, with the moving average calculation period along the left (rows) and the stop-loss values at the top (columns). Figure 21.5 shows that the fewest number of days and the smallest stop-loss value will give the score in the top left corner box. The bottom right box will show the result of the longest trend and largest stop-loss. With this presentation, there is continuity of performance in all directions. The patterns in the upper right may be similar to those of the lower left if the system displays some symmetry.

Although the combination of markets, trading rules, and parameter values can produce a very large combination of results, three patterns are most likely to appear in this 2-dimensional presentation. Figure 21.6a shows the simplest case, where a single area of successful performance is broad and gradually tapers off. This is analogous to the single parameter test shown in Figure 21.4a. Selecting the parameters associated with the center of the best performance area seems the most logical based on the reasoning that moderate shifts in price patterns are still likely to be profitable.

Two areas of profitability are usually the results of a highly volatile market (Figure 21.6b). Prices that are moving quickly and sustain a major trend can be traded using a fast or slow model. The fast model is often more profitable because it captures more of the trend; it then reacts faster to change of direction. It also has more losing trades and higher cost. The slower trend-following approach captures less of the price move by entering later and exiting later, but has a higher percentage of profitable trades. It may also have a lower information ratio because it holds those trades through larger
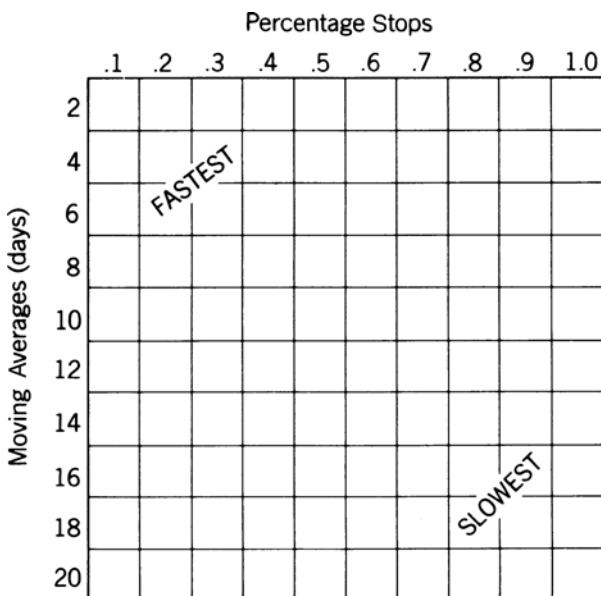
FIGURE 21.5 Standard configuration for a 2-dimensional optimization.
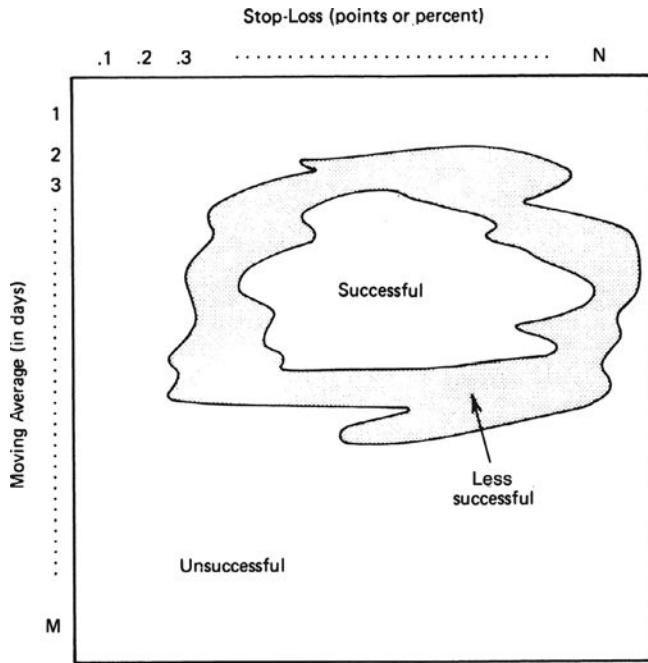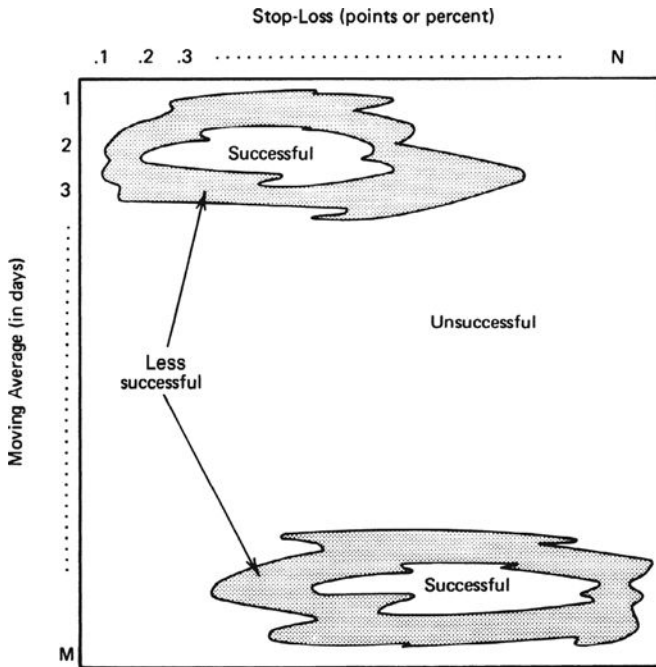
FIGURE 21.6 Patterns resulting from a 2-dimensional optimization. (a) Single profitable area. (b) Two distinct profitable areas. (c) No obvious pattern.
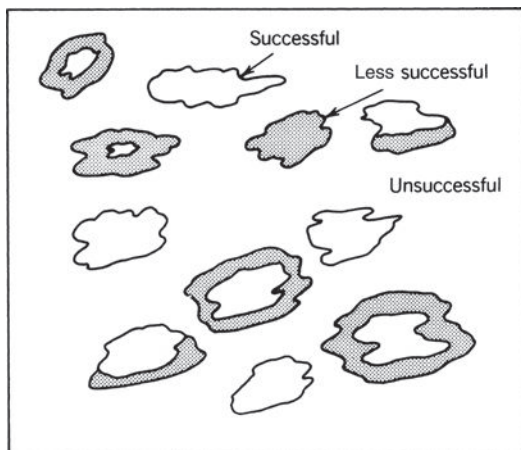
(c)

FIGURE 21.6   *(Continued)*

price swings. Choosing the parameters to use in actual trading follows the same reasoning applied to Figures 21.4b and 21.4c.

The third case, shown in Figure 21.6c, has erratic returns. The absence of a consistent pattern in the performance indicates that the trading strategy does not work with this data.

Alternate Ways of Visualizing Results

A wide range of statistical software provides user-friendly ways to show results as a *contour* or *topological relief map*. The patterns shown in Figure 21.6 can be seen in multiple colors, graded from one shade to another as returns increase and decrease, making it much easier to see the peaks and valleys of performance and get a sense of smoothness. Any spreadsheet program will allow a *surface plot* which will give you a rough look at the continuity of a 2-parameter test, as shown in Figure 21.7. This 3-dimensional graph gives the profit factors for the moving average–stop-loss strategy shown in Table 21.1. The moving average calculation periods are shown along the bottom right scale (Y), the stop-loss percentages along the bottom left (X), and the profit factor on the upright axis (Z). The returns shown on the surface have the best returns as the moving average calculation periods get larger, and the best stop-loss is in the center, about 2.5%. The fastest trends and the largest stop-losses have the worst performance. Overall, the stop-loss choices do not contribute as much as the trend. If smaller increments were used for testing the variables, the surface diagram would be smoother.

Visualizing with Scatter Diagrams
Unfortunately, technology does not yet allow us to chart more than 3 dimensions. If the system has three parameters, a moving average, RSI, and stop-loss, it will be necessary
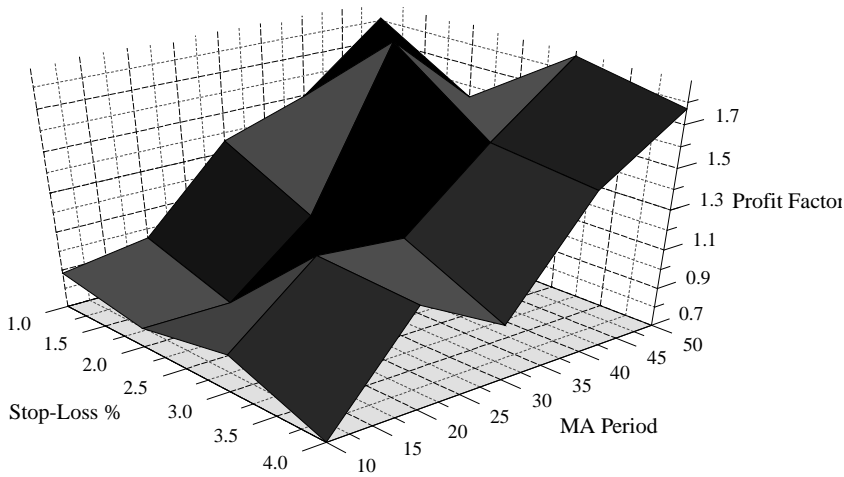
FIGURE 21.7   3-Dimensional surface plot of moving average–stop–loss strategy.

to have more than one chart in order to visualize the results. In Figures 21.8a and b, there are only two parameters, a moving average and a stop-loss. Each chart will show the results of only one. To accomplish this, the results of an optimization were exported to PSI-Plot software,[5] although a similar result could be achieved with Excel using an *XY plot*. Each circle represents the results of an individual test where the moving average or the stop-loss had the value shown along the bottom.

The moving average results in Figure 21.8a show a clear progression of success as the calculation period increases, except for the last two cases, the 40- and 50-day periods. While the range of results is the same, the 50-day clusters near the top, and the 40-day near the bottom, favoring the 50-day. Overall, we see that long-term moving averages perform better than shorter-term. Of course, this is only one market, one time period, and a very small test, so much more testing would need to be done to be confident of this conclusion.

Figure 21.8b shows that the stop-loss results are far less convincing. The 4% stop can be eliminated as the worst performer because it has lower highs and lower lows; however, the other three are not as clear. The one good note is that about 80% of the tests are above the profit factor value of 1.0, which indicates a net profit (gross profits greater than gross losses). While the 1% stop has the smallest range of results, the 2% the largest, and the 3% in between, it is not clear which is best. We can suggest that, if the averages are the same, the stop-loss value with the smallest standard deviation of profit-factors would show the greatest stability.

---

[5] *PSI-Plot* and *ProStat* are products of Poly Software International. They provide excellent statistical analysis and graphics at a reasonable cost.
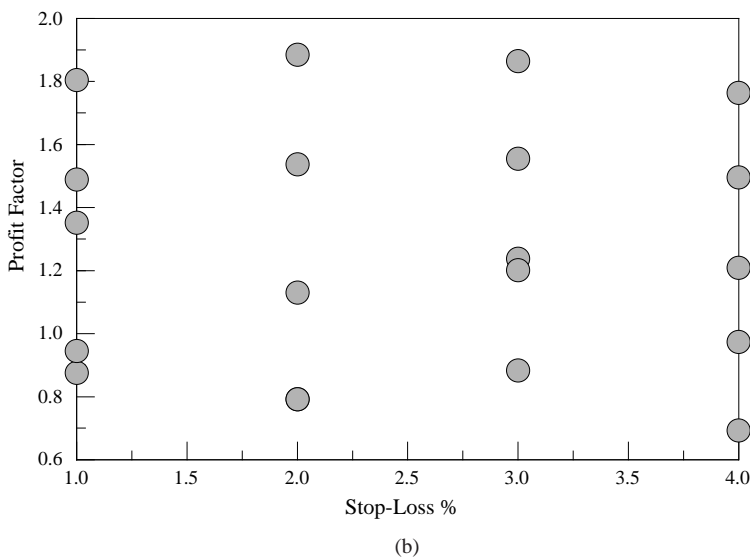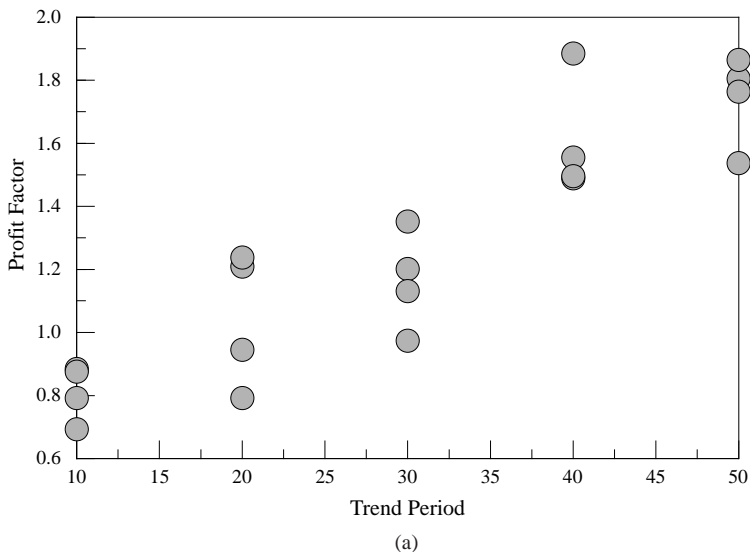
FIGURE 21.8   Visualizing the performance of multiple parameters.
(a) Moving average calculation periods. (b) Stop-loss % values.

Standardizing Test Results

When testing a single system, the results of one set of parameters can easily be compared to other sets because the test period, commissions, and other basic values are all the same. However, over time you will test many different systems and variations of those systems, and you will want to compare them to decide which strategy is best. This

requires some advanced planning. Standardizing test results is the best way to increase the usefulness of extensive testing. This should include:

1. *Annualizing all values.* Tests not performed using the exact same test intervals (for example, 10-year and 12-year periods), or tests on different futures markets, are likely to use data that starts and ends on different dates. Annualizing the results will not give a perfect comparison, but will make comparisons more relevant.

2. *Risk adjusting.* A profit of $10,000 with a drawdown of $5,000 is effectively the same as a profit of $20,000 and a drawdown of $10,000, given adequate reserves. Expressing the results as profits divided by risk, called the *information ratio*, is a better way to view returns.

3. *Adjusting for standard error.* The important difference between two tests that yield the same results, one test with two trades and the other with 50 trades, is that the one with only two trades is a less reliable result. If both posted 50% profitable trades, another losing trade would make the reliability of the second test drop to 49%, while the first case would drop to 33%. Instead, results should be presented at the same confidence level, which means subtracting the standard error from the current result.

Calculations for these three measures can be found in Chapter 2.

Averaging the Results

Because the testing process is computer-dependent, it is desirable to make the selection of the best parameters automatically. This can usually be done effectively by averaging either the annualized returns or the return ratio displayed in the test map. In the case of a moving average system, you may choose to eliminate distortions in performance caused by spikes. For example, the results of a 3-, 4-, 5-, and 6-day test show profits of 1,000, 8,000, 3,000, and 4,000, respectively. The 4-day case appears to be a profit spike and could be minimized by replacing its value with the average of the three values or, preferably, by the interpolated value of the two adjacent points (Figure 21.9). In general, the substitution process in which every value is replaced by the average of its neighbors

$$\overline{PL}_i = \frac{PL_{i-1} + PL_i + PL_{i+1}}{3}$$

will smooth out the returns. The best parameter set is the highest value of $\overline{PL}_i$ once all averages have been substituted for the raw test results. This may help find the overall performance pattern rather than an outlier.

2-Parameter Averaging

The results of a 2-parameter test may also be averaged to create a smoother set of values. Using a 2-dimensional grid shown in Figure 21.10, the original net profit or loss of each test will be denoted as PL_{ij}, associated with row i and column j of the display. The object is to replace each PL_{ij} with $\overline{PL}_{ij}$, the average value of its neighbors. This is done,
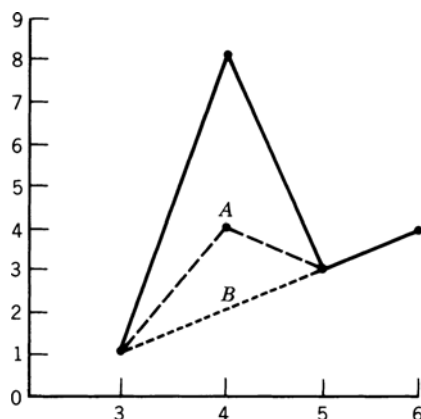
FIGURE 21.9   Replacing a spike with the average A or an interpolation B.

as seen in Figure 21.10a, by taking an average of the 8 test results adjacent to the ijth value as well as the center value. There are special cases when the ijth result is not fully surrounded but on the perimeter of the test grid. Figures 21.10b, c, and d show the averaging technique used when the ijth box is a top, side, or corner value. The 9-box average shown here is comparable to the 3-point average in the 1-parameter test. If the test has relatively small increments between tests and more test cases, an average of a larger area would be better. A generalization of this method can be found later in this chapter under "Sensitivity Testing."

LARGE-SCALE TESTING

A strategy with many parameters, each with a wide range of values, will result in a large-scale test—an optimization of hundreds of thousands (or even millions) of tests. It would be impossible to use any form of sequential testing to reduce the test time to something reasonable. Yet there are search methods that can arrive at a likely solution quickly. Two of these are *genetic algorithms* and *Monte Carlo* sampling.

Genetic Algorithms

The use of genetic algorithms for finding parameters was discussed in Chapter 20, in the section "Genetic Algorithms." A genetic algorithm is a very practical and very powerful searching tool.[6] It would be best to read the sections in Chapter 20 before continuing here.

---

[6]Readers are encouraged to read Murray A. Ruggiero, Jr., "Using Genetic Algorithms for Trading Applications," Chapter 20 in *Cybernetic Trading Strategies* (New York: John Wiley & Sons, 1997).
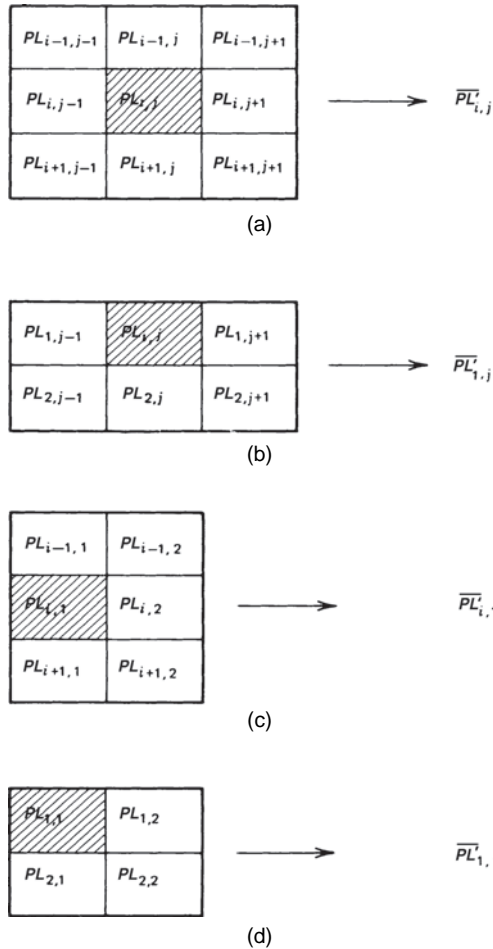
FIGURE 21.10   Averaging of map output results. (a) Center average (9 box). (b) Top edge average (6 box). (c) Left edge average (6 box). (d) Upper left corner average (4 box).

In the previous discussion, a trend and an oscillator were combined into a trading strategy and the rules for applying each component could be randomly changed. In this example, the rules will be fixed, but the parameter values are unknown; only the range of parameter values can be specified. Define the strategy and the range of calculation values as:

1. A long-term moving average trend, 30 to 250 days, increments of 5.

2. A percentage band around the long-term moving average, 0.5% to 4.0%, increments of 0.25%.

3. A short-term moving average trend, 3 to 30 days, increments of 3.

4. A percentage band around the short-term moving average, 0.1% to 1.5%, increments of 0.1%.

5. An RSI for entry timing, 5 to 20 days, increments of 1.

6. RSI overbought and oversold thresholds, 40 to 95 for overbought, −40 to −95 for oversold (only one set is needed if these values are symmetric), increments of 5.

7. A percentage stop-loss, 0.25% to 3.0%, increments of 0.25%.

8. A moving average trend for equity, 5 to 50 days, increments of 5.

These 8 parameters seem innocent enough, but total 2,332,800,000 tests. At 10 tests per second, the processing time would be 64,800 hours, more than 7 years! Following the steps discussed in Chapter 20,

Initialization
1. A *string* of 8 cells is created to represent each of the parameters. A string is the general term given to a single set of parameter values, called a *chromosome*.

2. Each cell in the string is assigned a random number that determines which parameter value will be used.

3. A pool of 1,000 to 5,000 strings is created, the larger, the better.

Propagation
4. The *objective function* (ranking score such as the information ratio) is calculated for each string.

5. The pool is sorted in descending order by objective function.

6. Multiple copies are made of the highest-ranking strings and added to a new pool.

7. The new pool is sorted by highest objective function, and culled to the best 5,000 strings. The new pool replaces the old pool so there is only one pool.

Mating
8. Two strings, not previously selected, are chosen at random from the pool.

9. A random number between 1 and 8 divides the cells of the strings into left and right parts.

10. The left part of the first string is switched with the left part of the second string, creating two new strings.

11. The mating process continues until 10% to 15% of the strings have been combined.

Mutating
12. One string is chosen at random from the pool.

13. One of the 8 cells is chosen at random.

14. Based on the range of values for the parameter assigned to this cell, a new value is chosen at random and replaces the old value.

15. Mutation continues until one cell in 10% of the strings has been changed.

Converging

16. The best objective function, or score, is saved at the end of each pass, and the parameter values associated with this string are saved.

17. The genetic algorithm process continues at Step 4 until the objective function converges. That occurs when, after a number of passes, the objective function does not increase significantly.

18. The genetic algorithm process is repeated up to 5 times, beginning with the initialization, in order to produce different seeded strings.

19. After 5 passes, the string previously saved with the highest objective function is the solution.

Saving a set of best results will improve your understanding of its success. If, for example, the best 20 strings all cluster at a similar solution, it is likely to be more robust. If the cells of the top 20 vary considerably, then the integrity of the results are in question.

For a trading strategy with 8 parameters and over 2 billion combinations, the genetic algorithm should converge within 100 passes. Repeating the process five times in order to seed the cells differently increases the number of passes to 500. It is likely that the elapsed computer time will be 30 minutes. It does not guarantee that the maximum combination will be found. The chances of finding a maximum can be increased with a larger pool and by reseeding more times. However, the parameter values chosen from this short search will be astoundingly good. We will never know how close the results come to finding the maximum because no one is going to run a full test. A more detailed example of the use of a genetic algorithm, applied to portfolio allocation, can be found in Chapter 24.

Monte Carlo Sampling

The problem with the genetic algorithm solution is that it returns only the maximum results. There is no information on the depth of success that is needed to determine robustness. It is helpful to know the parameter values that are most successful, but it is not always enough. *Monte Carlo* sampling is used to create a statistically valid subset of tests when the total number of tests is too large. The results of this smaller set of tests can be analyzed as though it were the complete set.

Monte Carlo simulation (MCS) was created by the famous mathematicians Johann von Neumann and Stanislaw M. Ulam; they devised the method shortly after World War II to help them simulate the behavior of an atomic weapon. There is an interesting example of how the Monte Carlo process has been used to find the area of a lake with an

extremely irregular shoreline. A picture of the lake is taken from above and the scale is determined (e.g. $1'' = 100$ feet). Think about the picture as $10'' \times 10''$, displayed in color on a computer screen, with the lake clearly blue and the surroundings green. The picture is then superimposed with a fine grid, so that there are 100 (or even 1,000 if more accuracy is needed) pixels per inch, for a dimension of $1,000 \times 1,000$ representing the entire screen, an unknown part of which is the lake. Now pick two *uniform random numbers* between 1 and 1,000 to represent the x and y coordinates, and say the first values are 135 and 441. Remember that a uniform random number has an *equal chance* of being any of the values between 1 and 1,000. The position (135,441) is marked on the display and the computer determines whether it falls in the lake or on the shore. This process is repeated for 1,000 pairs of random numbers. When completed, the computer has counted 520 pairs that were in the lake; therefore, 52% of the display area is lake, or 520,000 square feet. If more accuracy was needed, the grid would be divided into more pixels, and more random pairs would be generated. The process works because the probability of a pair of random numbers falling in the lake is equal to the ratio of the lake's area to the entire picture.[7]

Random numbers can also be used to generate parameter values in a test of a trading strategy. Using the same 8 parameters applied to the genetic algorithm solution, a value is picked at random for each parameter from the list of ranges specific to that parameter. This is the same step as the initialization of the pool of strings used in the genetic algorithm. Every time a value is picked a counter is incremented. Every parameter has a counter for every possible value used in the test. After the first string is created and saved in a table along with its performance statistics, the counter for each random value that was used is increased by 1. After 1,000 random strings, the total count for the list of values in any parameter will total 1,000.

The first opportunity to end the test is when the distribution of all values within a parameter list is normal. For a truly random selection, each value should have approximately the same count. Because there were 10 values for parameter 8, after 1,000 strings each value should have a count of very close to 100. If there were 20 values, each would have a count of 50. Once the minimum distribution criterion is satisfied, the more strings that are generated, the better the sample. These strings can now be analyzed as though they were the entire test set, although academics will argue that the statistical quality of the sample data (e.g., variance) may vary significantly from the whole set.

As with the genetic algorithm solution, the Monte Carlo sampling can be performed a number of times. If the best results are similar, then you can be confident that the smaller sample is a good representation of the entire set. The use of random values has the additional advantage of avoiding unfair targeting of values that may be known in advance to perform well.

---

[7] Adapted from an example in Wolf von Ronik, "System Testing via Monte Carlo," *Futures* (April 2001).

REFINING THE STRATEGY RULES

As testing progresses, still using the in-sample data, it is inevitable that the analyst will want to modify a rule or add a new feature to the trading strategy. This usually follows from inspecting the results of tests and noting that a specific pattern was not treated properly (or profitably). After some work, the analyst introduces a carefully constructed rule, which turns a previous loss into a profit. Figure 21.11 shows two possible changes in the performance pattern for 1-parameter test based on this change.

If the rule change improves the specific situation that was intended, but loses money in other market situations, the complete test pattern will appear as shown in Figure 21.11a. Higher peak profits occur where the new rule works best, but other parameter values have been turned from profits to losses or to larger losses. The average returns may remain the same, but the results are more extreme. Because the new rule caused the fitting of a specific pattern at the cost of added losses in other patterns, it is not an improvement. Statistically, this performance change can be measured by kurtosis. A very high value of kurtosis is a sign of overfitting.

In Figure 21.11b, the new rule improves performance in all cases. This pattern is an ideal solution. It is possible, of course, that an improvement is defined in such a way that it corrects only that specific problem without affecting any other trades. That type of rule
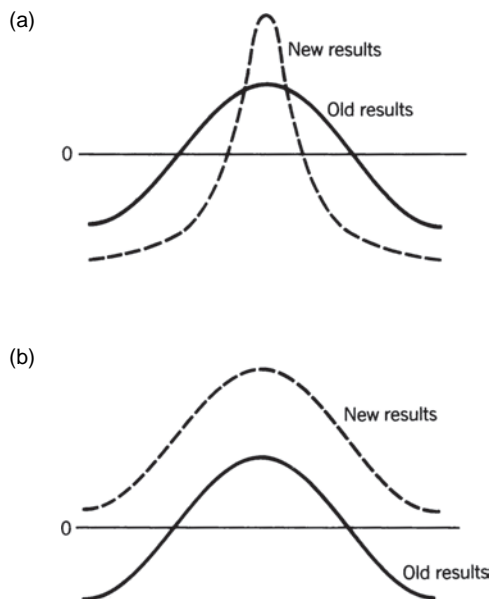
FIGURE 21.11 Patterns resulting from changing rules. (a) A rule change that improves one situation at the cost of others. (b) A rule change resulting in general improvement.

fitting, which affects only one or two cases, is unlikely to work in the future. If it falsely raises expectations of success, it puts the investor at risk.

ARRIVING AT VALID TEST RESULTS

It is not unusual for the results of an optimization, especially one with many parameters, to appear perfect. All the trades could be profitable, the equity drops might be small, and the strategy might perform in changing markets, and yet, the final system will consistently lose money when traded. Creating a system with predictive qualities (rather than one that is historically fitted) requires preparation in advance of testing.

There is no "best" place to say this, and it should be said in more than one place. *We have a tendency to accept good test results, but look closely for errors when there are bad results.* It is important that all results be verified carefully.

Common Sense versus Statistics

You can always find a particular moving average or indicator that gave a signal to sell the S&P just before the 1987 crash, or ahead of the 2008 mortgage crisis. It is not difficult when a computer can combine millions of combinations of trend speeds, filters, momentum indicators, divergence rules, and so on, using powerful search engines. It is also easy when you are willing to sacrifice performance at other times to capture the biggest profits in recent history. However, what works in a specific case, and often an unusual one, is not likely to work in general. Manipulating the rules is a wasted effort. Showing that a combination of technical indicators would have profited from the last market crash has no bearing on how it will respond to the next crisis.

The adage that "there are lies, half lies, and statistics" is true. Statistics can prove a point or prove just the opposite, depending on how the numbers are presented. A frequent contradiction in the news is when an anticrime group states that burglaries doubled in the past year while the police department states that burglaries rose by the slowest rate in 5 years. A market can be in an uptrend and a downtrend at the same time, based on the time interval over which you view it.

The most profitable test results may not be the ones you want. If you tested the S&P from 1990 through 2003, the results would be overwhelmed by the 10-year bull market. The parameter values may be biased toward longer-term trends that captured most of the market's upwards move; however, it would have lost money during the sideways prices moves of 2002–2003. The statistical profile will still look good because the bull market profits were far greater than the more recent losses; however, using a strategy that capitalized on an unusual bull market is not a good choice. As Jack Schwager has stated, "The mistake is extrapolating probable future performance on the basis of an isolated and well-chosen example from the past."[8]

---

[8]Jack D. Schwager, *Schwager on Futures: Technical Analysis* (New York: John Wiley & Sons, 1996), 673.

Omissions and Scenarios

The more difficult problems to overcome are the ones you cannot see, the price moves that are not part of your test data, and the *omissions*. Every few years the market experiences a price shock, most recently the subprime collapse of 2008, before that the terrorist attack of September 11, 2001. Price history is full of small price shocks caused by economic reports, but the very large ones are years apart. If your test does not include data with a significant price shock, then you cannot know if you will survive the next one. If you have tested a strategy using data that is highly skewed in one direction, as the equity markets were in the 1990s, then the resulting strategy will be biased. You must be constantly looking for potential problems. As of early 2012, the best strategy would have been long bonds from the early 1990s, ignoring all information to the contrary. But eventually inflation and an improved economy will change that, and a long-only strategy will give back its gains. You must question the ability of your strategy to perform under different scenarios. The following guidelines were developed over many years and are intended to avoid the most basic and costly errors that occur in testing. You may be able to add your own experiences to this list.

Searching for Robustness

A *robust* system is one that performs consistently in a wide variety of situations, including patterns that are still unforeseen. From a practical view, this translates into a method that succeeds with the fewest parameters tested over the most data. In tests by Futures Truth,[9] the best performing systems commonly have four or fewer variables.[10] The characteristics of a system with the best forecasting ability are:

1. *It must be based on a sound premise.* Each rule and formula must capitalize on a real fundamental or price phenomenon. *Discovering* a price pattern or cycle through optimization may seem to be a revelation, but it is more likely to be an illusion. By testing enough patterns, it is statistically probable that one of them will seem to fit. Without a fundamental reason for the existence of that pattern, it is not safe to use it. For example, the use of a very long-term trend to trade interest rates is intended to mimic the direction of U.S. economic policy; a shorter 1-month trend may be best for capturing the seasonal moves in agricultural commodities. A short-term trader may hypothesize that a strong close on increasing intraday volatility will have follow-through on the next open. Observations based on experience and confirmed by testing are usually sound.

2. *It must adapt to changing market conditions.* A system that assumes that markets do not change or that everything has been seen in past data will suffer

---

[9]John Hill and George Pruitt, Futures Truth, Hendersonville, N.C., is an organization that validates and monitors trading systems. It publishes *Futures Truth Magazine*, www.futurestruth.com.

[10]John Hill, "Simple vs. Complex," *Futures* (March 1996), 57.

large equity swings. Self-adjusting features do not use absolute price levels or fixed dollar amounts. They might include an inflator or deflator, stop-loss values that adjust to volatility, rule shifts based on whether the current year shows a seasonal or nonseasonal pattern, and a variation in position size based on changing risk.

3. *It must be tested properly.* The principles of statistics say that the best tests use the most data. More data includes bull, bear, and sideways markets, large and small price shocks, and periods of instability and doldrums. There is no substitute for more data, even when some analysts say that old data is no longer valid. Proper test procedures include using in-sample and out-of-sample data periods, discussed earlier. Alternately, you can paper trade until there are enough trades to compare the out-of-sample profile with the expectations defined by your tests.

From time to time, each of the three points above will seem to be the most important. In the long run, they must all be included to create a robust program.

Performance Criteria

Assuming that the proper principles have been followed, are the test results good enough? Although the strategy shows steady profits with a low equity drop, could a simpler system have done better? It can be a rude awakening to be pleased with your system's 20% returns only to realize that the average stock return was 25%. The industry has become highly competitive.

Benchmarks

It is necessary to have a *benchmark* that provides a way of measuring success. It is best if this is a well-documented indicator, such as the S&P Index, the Lehman Brothers Treasury Index, a class of funds, such as Fidelity, or a Commodity Trading Advisor (CTA) Index as published by BarclayHedge. Benchmarks add a note of reality to expectations. The reason for developing a trading strategy is to have more consistent profits, or less risk during volatile periods, or a different return profile, than an investment program already offered by, say, the Royal Bank of Scotland. Depending on the period you measure, the S&P 500 has a return-to-risk ratio ranging from about 1.0 to zero, meaning that, at the better end, there is a 16% chance that the risk during any one year will exceed the annualized returns.

The spectacular profits or losses of one manager that get spotlighted in the news is not a good measure of anything, but simply a rare event. The focus should be on consistency, having very good gains each year. The largest gain can only come with high risk. You should not be surprised that those investment advisors posting the highest monthly returns have had very erratic performance in the past. It is all too common to see a fund with the largest monthly gains, up 20% in June, but still down 5% for the year.

Inaccuracies can hide in the benchmarks. Two important idiosyncrasies of published benchmarks are *survivorship bias* and *asymmetric measurement of returns.* Considering the last, some hedge funds that lost 50% in 2008 then gained 50% in 2009. Did they recover their investors' losses? Of course not. A gain of 50% after a loss of 50% still leaves the investor with a loss of 25%. Survivorship bias is a different type of problem. If a hedge fund has been reporting its performance each month, but has a fatal loss and closes down its fund, it most often will not report the returns of the last month. It has no motivation to do that. Therefore, the performance benchmarks often omit the largest of the negative monthly returns.

Measuring Test Results

As a preamble to measuring results, specify the test range for each parameter in advance. An optimization should represent a broad sampling of tests over a wide but reasonable range of parameter values. This range should be established before testing begins and not changed. For example, if a short-term moving average is being tested from 3 to 15 days, then an RSI used for timing would have a range of 3 to 10 days. You would not test the RSI from 3 to 50 days because the longer periods do not make sense in the context of the original plan. Even worse, if the 45-day RSI proved the best, the initial concept may be discarded in order to pursue a solution that is most likely to be an overfitted dead end. If the common-sense ranges are used and are not profitable, the concept should be reviewed.

There are a number of popular performance criteria used to evaluate any trading strategy.

1. *Information ratio.* This single, most important measure is the annualized returns divided by the annualized standard deviation of returns. It gives you a clear way to value the return relative to risk and the smoothness of returns. It is discussed in detail in Chapter 23, under "Measuring Return and Risk." Other measures, such as TradeStation's *profit factor*, the gross profits divided by the gross losses, can also be used.

2. *Net profits or losses.* Although not the best criterion for picking a parameter value, profits are the reason for trading. While no one would select a system that produces a net loss, it is the other statistics that will tell which are the best choices for parameter values.

3. *Number of trades.* This simple value indicates whether your test was long enough to generate dependable results. A few trades can appear very profitable, but small numbers are statistically unreliable. A system that has more trades with the same returns has a better chance of performing up to expectations.

4. *Percentage of profitable trades.* Also called *reliability*, a value above 60% can be interpreted as a method that captures profits regularly. A trend system will be working correctly if its reliability is between 30% and 40%.

5. *Average net return per trade.* The average return per trade gives you an indication of how difficult it will be to realize the system returns. An average return of $50 per

trade for a currency is likely to be overwhelmed by the slippage, which typically exceeds 4 bps for each side, a total of $100 for a round-turn futures contract. Many very fast trading systems look good until you apply realistic costs. Slippage per side on the S&P should be 50 basis points, fixed income 1/32 to 2/32 (about $45), and energy products up to 20–30 basis points (about $100). The slippage executing stocks varies with the liquidity and market volatility. Slippage should be as accurate as possible, determined from actual trading. While it is good to be conservative by using a larger slippage, it eliminates faster trading systems that may be highly profitable.

6. *Maximum drawdown.* The largest equity swing from peak to valley is the maximum drawdown. This measurement can be very erratic and is not likely to be the largest risk seen in the future; however, it gives a rough idea of the minimum capital needed to trade this market. A surprisingly small maximum drawdown is a sign of overfitting or a very small test period. It is safer to take the average maximum drawdown from a range of tests than the one from the best result. An investor will typically capitalize a trading account with three times the maximum drawdown.

7. *Annualized rate of return.* The rate of return requires that the size of the initial investment be known. The investment can be calculated after the fact from the risk-to-equity ratio based on the maximum drawdown or standard deviation of historic returns, linked to a *target volatility level.* Annualized returns allow you to compare one test against others or against a benchmark.

8. *Average time-to-recovery.* A large drawdown may be inevitable in a realistic system, but a shorter time-to-recovery is most desirable. *Time-to-recovery* is the time between two successive equity highs. A larger drawdown with a much faster recovery seems to be a better trade-off for most investors.

9. *Time in the market.* All else being equal—profitability and risk—a trading system that is in the market less than another system is preferable. If two systems have approximately the same performance profiles, then the one that is in the market less is exposed to fewer price shocks. It may be a good trade-off to sacrifice some potential profit in order to be out of the market more often. As discussed in Chapter 22, being out of the market is the only defense against event risk.

10. *Slope of periodic returns.* If the quarterly or annualized returns are saved and a linear regression line drawn through those returns, the slope of that line would indicate whether the performance of the strategy is constant over time. A declining slope should be interpreted as better performance in the past, perhaps a trend system that did exceptionally well in the equity market of the 1990s but is not doing well now. A rising slope means that the strategy does better under current price patterns than older ones. A horizontal slope is best, indicating consistent performance across all time periods.

Having decided the measurements needed to understand the results of an individual test, there are other rules that must be established, including how to assess the end results. This will help to avoid bending the rules to fit the results.

Average Results

Use the average of all test results. Specifying reasonable parameter ranges is important when evaluating the test results. Nearly all sets of tests will show losses, but hopefully, there will be some areas of attractive profits. If you tested 1,000 cases and 30% of the tests showed returns of about 25% per year, 30% showed breakeven results, and the last 40% showed various profits and losses, you might say that the 30% profitable tests are a broad area from which parameter values can be chosen. That assumes that the market will continue to perform in a way that allows those parameters to generate profits during the next year. It is better to assume that the price patterns change; you cannot tell which combination of parameters will be the best. Regardless of the past returns for the parameters you choose, your expectations should be the average performance of all tests.

It would be optimistic to expect the average return of all tests to be highly profitable; however, that is the correct goal. When comparing systems, the best one has the highest average of all tests as well as the most profitable number of tests. If you accept the premise that actual trading performance is represented by the average of all tests, then your expectations are realistic. A statistical indication of the likely success of a strategy would be

Adjusted returns = Average of all tests – 1 Standard deviation of all test returns

A system satisfying this requirement should have an 84% chance of achieving the adjusted returns.

Review Outliers

Realistically, a system should not base its profitability on a single major market move over the test interval. In particular, if the move that generated the profit was actually a price shock, the system had only a 50% chance of posting that profit and has an equal chance of a loss in the future. It may be practical to remove the profits generated on the first day or two of a move that began with a price shock, allowing for the possibility of holding the wrong position in the future. By posting a windfall profit, the risk of the system is erroneously reduced. Price shocks will be discussed in the next chapter.

Concentrate on large profits and large losses

The largest profits and losses are critical to the net performance of any system; therefore, you should make a special effort to study those trades. In doing that, you may find that the largest profits were the results of price shocks, as discussed in the previous point, or that there is an erroneous piece of data that was not seen before. If there are a number of large profits and no large losses, you can be sure that you have over fit the data.

Sensitivity Testing

Once the parameter values have been selected, a sensitivity test should be performed. *Sensitivity testing* is simply a way of finding out how much performance changes when

the individual and combinations of parameter values are shifted up and down by small amounts. A robust solution should be modestly sensitive to changing values; in some cases, the performance will improve, but generally, as larger changes in values are used, some decay is expected. Overall, small changes in parameter values should not jeopardize the profitability of the strategy. If the performance declines for all other parameter values, then the current model is a peak solution, which usually means it was overfit and will not meet expectations.

Sensitivity in *n*-Space

When using multiple parameters, sensitivity needs to be viewed as an n-dimensional problem, commonly called *n-space*. Let's try to visualize what this means and find a way to show sensitivity given changes in each of n parameters.

Using four parameters in our example, assign w_0, x_0, y_0, and z_0 as the chosen, or *core*, parameter values with the resulting performance at location (0,0,0,0). As the parameter values change, the new results will be away from the core in any of 4 directions. With n parameters they would move in n directions. We can measure where they are by finding *the radius of an n-dimensional "sphere,"*

$$r_i^2 = (w_i - w_0)^2 + (x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2$$

Because the position of the test is defined by the radius, the distinction between these tests will be how far they are from the core. All i sets of parameters can be sorted, smallest to largest. If they are partitioned into, for example, 10 equal sets, they represent concentric layers of the same sphere going outward. By averaging the performance statistics of each of the tests within the same layer, we can see how much that average decays as it moves away from the core in any direction. We then have an n-dimensional sensitivity test. An enterprising analyst might take this a step farther and use random numbers for the parameters (within a predefined range), calculate the radius from the midpoint parameter values, along with the performance statistics, and find the most robust layer.

Reading between the Lines

It is easy to show specific examples of rules, indicators, and systems that produce large profits, while the overall performance is bad. Many books and systems for sale base the "proof" of a trading method on specific examples. These examples look reasonable in isolation. There is an inclination to accept these examples if the concept is sound, when there is fundamental or economic substantiation, such as expecting a seasonal low in the U.S. crops at harvest. Nevertheless, proving that a system is likely to be successful can only be done with a long-term test and disclosure of the statistical performance profile, not with a few well-chosen examples.

The Significance of "Significant"

It is often the case that *statistically significant* results are not actually significant. This goes back to the problem of testing, retesting, and knowing how many combinations have been included in the sample. Using a less-than-statistical method, we previously suggested that a trading method was robust if 70% of all tests were profitable. Naturally, that would be initial tests using a reasonable range of test values. If we subsequently narrowed that range to eliminate calculation periods that always lost money, that rule would no longer apply.

If 1,000 different parameter values, spread over a wide range, were tested on a 10-year period of historic data, how many of them need to be profitable to consider the system statistically successful? That's not clear. But, given the number of tests, one highly successful parameter is certainly not significant. In addition, if you continue to try new trading ideas and one is finally successful, is that really significant? Isn't it the cumulative number of tests across all strategies that should be considered? Finding 1 strategy in 100 does not seem as comforting as finding the first one successful based on a sound premise. Cranking out combinations of indicators and trends is sure to find an occasional but insignificant success.

Systems That Work in Only One Market

If a system is developed for the *e*mini S&P futures, should it work on the NASDAQ 100, or on the Dow? Seen on a quote machine, and statistically, the price moves of those markets are very similar. When an Initial Claims report is surprisingly high, they all drop together; and when Retail Sales are lower than expected, they all fall together. Each of the three markets represents a different set of stocks, and each has a different volatility profile, and while a disappointing earnings report by Intel will impact NASDAQ more than the S&P or Dow, there are implications in the Intel earnings that inevitably affect stocks in all markets.

Then the answer should be *yes*, a strategy that is profitable for one index market should be profitable for similar index markets, unless it is using certain characteristics unique to one class of market. While profitability may vary, if it generates good profits for the S&P and losses for NASDAQ, then the parameter values are most likely overfit. The two other index markets serve nicely as out-of-sample data.

Advertised Systems for One Market

From time to time, all traders receive mail that offers a highly specialized system called "The Cattle Trader," "The Gold Day-Trading System," or "Easy Profits through Stock Index Trading." It is best to assume that each of these systems has been finely tuned with rules unique to one market. Optimization is a well-known process and can be used responsibly or easily abused. Because these offers tend to have big profits with little risk, it seems reasonable that the proclaimed results are unrealistic. It is a clear case of "too good to be true." And, because the rules of the program are not revealed, it is impossible to test them yourself.

COMPARING THE RESULTS OF TWO SYSTEMS

The easiest way to understand the use of testing is by example. Using daily NASDAQ 100 data from January 1997 through December 2003, we compare the results of two trend systems, one based on a moving average and the other on the slope of a linear regression, in order to decide which of the two is best. These calculations can be found in Chapter 7. In both cases, the system generates a buy signal or a sell signal when the trendline or slope turns up or down, respectively. There are no other trading rules. Table 21.2 shows the results of varying the calculation interval from 10 to 250 days in steps of 10 days. Transaction costs of $125 per trade were used to cover commissions and slippage.

Which System Is Better?

The first point to notice is that most test results are profitable. This is seen in both the net profits (column 2) and the profit factor (column 3). The high percentage of good tests justifies the underlying premise—that the market has trends. We should always expect all basic trend systems—moving averages, regression, and swing—to be profitable if one is profitable because they all benefit from the same price moves, some more, some less. Figure 21.12 is a good example of the similarity. At the short end, both methods suffer from trading too quickly and paying a large commission. On the far right they track closely.

The next interesting statistic is the number of trades. This is a 6-year test, so if the number of trades is less than one per year, some investors may not consider it "active." In the *#Trds* columns for both systems, the number of trades declines in an orderly manner until it reaches 8, then it seems to stall. In addition, the profit factor (*PFact*) jumps up at the moving average period of 190 and the regression period of 140. This is exaggerated because TradeStation only counts closed-out profits, rather than the daily change in profits. Using an information ratio based on daily prices would be much more uniform.

We try to get a better picture of comparative performance by looking only at the moving average system through 180 and the regression slope through 130. The feature that ties them together will be the number of trades. It gives us an idea of the sensitivity of each trend indicator. You should notice that the number of trades declines faster with the regression method. Figures 21.13a and b compare the number of trades and the corresponding profit factor for each method. In part a, the moving average trades start higher (the left scale) and decline slower than the slope results in part b. While both methods now span the same ranges, the moving average results are more orderly.

Then which is the better trading system? The moving average method has the highest net profit at calculation period of 150 with 32 trades. The slope method peaks at 200 with only 2 trades. In our limited test range, it peaked at the longest calculation period of 130 with lower profits and few trades. Perhaps more important, the

TABLE 21.2 Comparison of NASDAQ Trend Tests

| | (a) Moving average | | | | | | (b) Linear regression slope | | | | |
Period	NetPrft	PFact	#Trds	%Prft	AvgTrd	Period	NetPrft	PFact	#Trds	%Prft	AvgTrd
10	−338005	0.71	229	28	−1476	10	−44455	0.95	173	34	−257
20	−70065	0.91	181	30	−387	20	49995	1.09	81	29	617
30	104715	1.17	157	34	667	30	267190	1.77	50	38	5344
40	226065	1.61	89	37	2540	40	109005	1.32	37	37	2946
50	198930	1.5	74	29	2688	50	130605	1.39	27	40	4837
60	284610	1.91	80	37	3558	60	157965	1.61	23	34	6868
70	261670	1.97	72	36	3634	70	300015	2.94	19	36	15790
80	−21600	0.95	86	26	−251	80	313145	3.39	15	40	20876
90	93940	1.25	78	34	1204	90	165265	1.85	15	33	11018
100	−9640	0.98	74	24	−130	100	190660	1.97	16	37	11916
110	62100	1.15	74	24	839	110	134610	1.55	12	33	11218
120	273470	2.17	50	38	5469	120	270560	3.17	12	41	22547
130	312970	2.73	36	33	8694	130	416630	6.96	8	62	52079
140	325790	2.57	48	27	6787	140	531740	16.53	6	66	88623
150	469530	5.52	32	40	14673	150	468600	14.72	6	66	78100
160	445130	5.23	38	34	11714	160	499400	10.11	6	66	83233
170	455190	8.58	20	35	22760	170	418800	9.63	6	66	69800
180	434260	6.17	18	27	24126	180	493850	27.05	4	75	123463
190	515210	36.35	10	50	51521	190	594060	100	2	100	297030
200	480260	76.04	8	50	60033	200	524160	100	2	100	262080
210	401590	10.01	10	40	40159	210	540260	100	2	100	270130
220	409960	13.19	8	62	51245	220	463710	100	2	100	231855
230	338160	5.29	18	55	18787	230	401260	100	2	100	200630
240	359210	9.18	10	50	35921	240	404310	100	2	100	202155
250	346560	16.99	8	37	43320	250	417560	100	2	100	208780
Average	254400	8.57	60	37	16324	Average	328756	32.32	21	61	91267
Avg 10–180	194948	2.62	80	32	5951	Avg 10–130	189322	2.30	38	38	12754
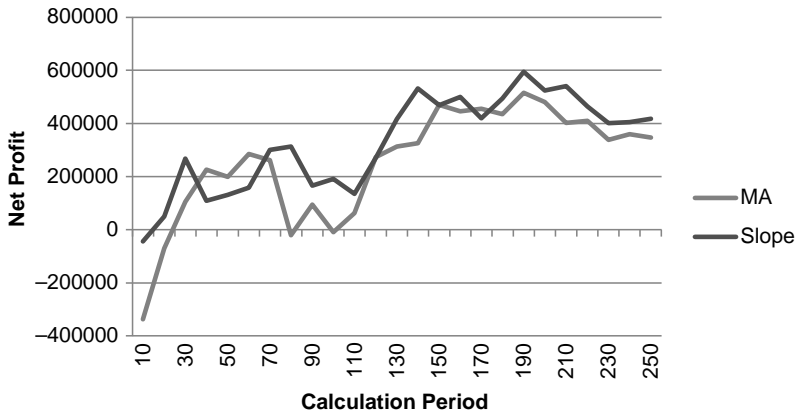
FIGURE 21.12   Net profit for moving average and slope methods, NASDAQ, 1997–2003.
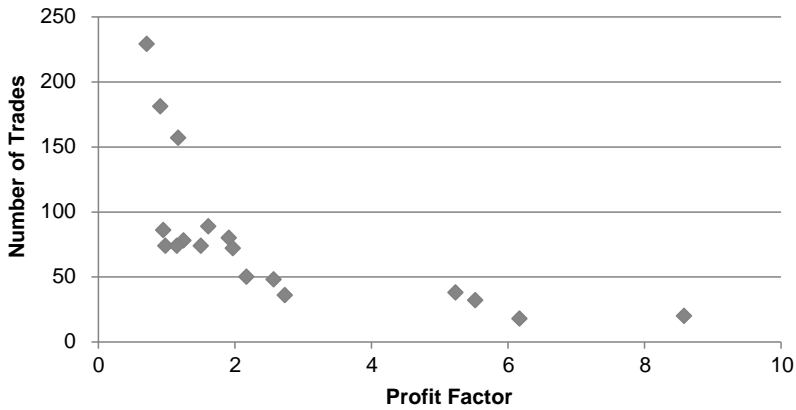
FIGURE 21.13(a)   Comparison of number of trades versus profit factor, moving average.
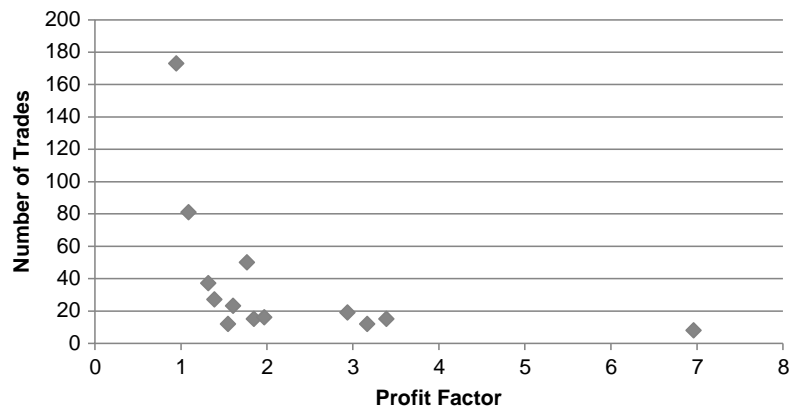
FIGURE 21.13(b)   Number of trades versus profit factor, regression slope.

average ratio for the moving average is 2.62 and for the slope 2.30 when we consider the smaller, more realistic ranges. *Because we cannot know which trend speed will be best in the future, the more robust system would be the one with the highest average ratio over given reasonable test ranges.* Check for performance decay as discussed earlier.

Using Correct Transaction Costs

Transaction costs have a significant impact on short-term trading. Faster trading has the advantages of smaller actual losses and more transactions. A system that performs well with more transactions will give more confidence of future performance. A system that only had two trades during the past 10 years may show high historic profits but also high risk and a poor sample.

On the negative side, short-term trading is a fight against slippage and overall transaction costs. There may be times when prices are moving so fast that you miss the trade completely. Most profits from short-term trading occur at the beginning of the trade and require good entry timing. Using a simple moving average method, this time for NASDAQ futures from 1991 through 2011, tests were run using no costs, then $50, $100, and $150 round-turn total costs. The results are shown in Figure 21.14. With no costs most of the tests that use more than a 15-day calculation period are profitable. As higher costs are added, the profitable part of the performance gets narrower, limited to calculation periods between 45 days and about 75 days. For many markets all tests above 45 days might be profitable; NASDAQ is not the best trending market.

The lesson in this simple exercise is that transaction costs must be included to be realistic, but costs that are too high will make everything seem to be a loss. Realistic costs are necessary.
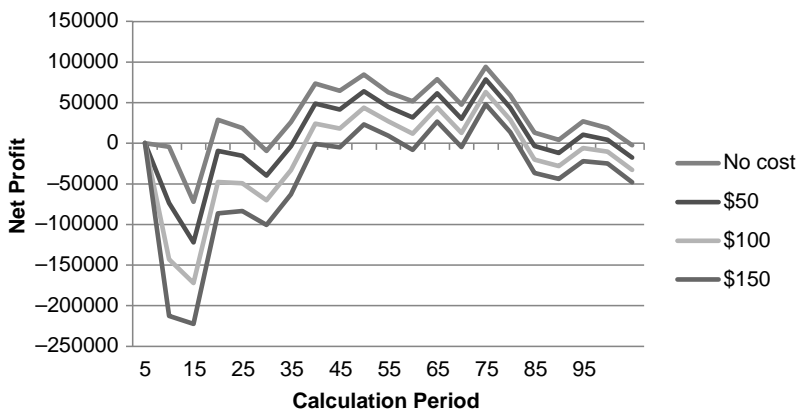
FIGURE 21.14 Comparison of net profits with varying transaction costs, applied to a moving average system on NASDAQ, 1991–2011.

PROFITING FROM THE WORST RESULTS

The worst results of an optimization or test sequence may hold valuable information. If the loss was the result of a single drawdown, then it is important to know how that drawdown was reflected in each of the tests. If a 30-day trend lost on a price shock but the 40-day trend profited, then the next time there is a price shock the opposite might occur. Price shocks are an extremely important concept and will be covered in the next chapter. Here it is only important to know that price shocks cannot be avoided, and generally cause losses. Choosing a parameter that side-stepped a price shock has no bearing on the future.

A persistent area of losing performance is different. Consider moving average results for NASDAQ from 1991 to 2011 all using faster calculation periods, shown in Figure 21.15. Any calculation period from 1 to 11 produced significant losses. From the previous discussions of trends and price noise, we know that nearly all equity index markets are noisy, and that noise generates losses in short-term trends. We can take advantage of this characteristic by using the short-term trend as entry timing for a longer-term trend that is profitable on its own.

If we consider this conceptually, we expect noise to cause losses for most trend systems using faster calculation periods. The results of a typical optimization of one variable can be seen in Figure 21.16a, very similar to Figure 21.15. When two moving averages are tested, there tend to be two clusters of returns, one showing losses where both averages have shorter calculation periods, and another with profits in the area of longer averages, seen in Figure 21.16b. An alternate strategy that *fades* the faster trend is,

> *Buy* when the slow average is up and the fast average turns down.
> *Sell short* when the slow average is down and the fast average turns up.
> *Exit the long position* when the slow average turns down.
> *Cover the short position* when the slow average turns up.
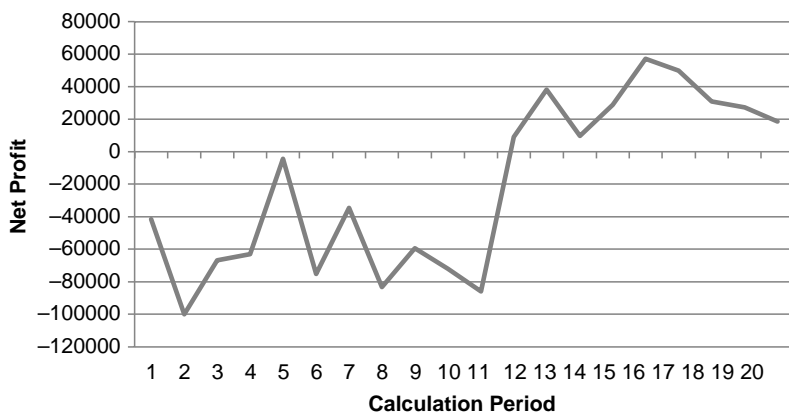
FIGURE 21.15   Net PL for a NASDAQ moving average system, 1991–2011, where the calculation periods are short.
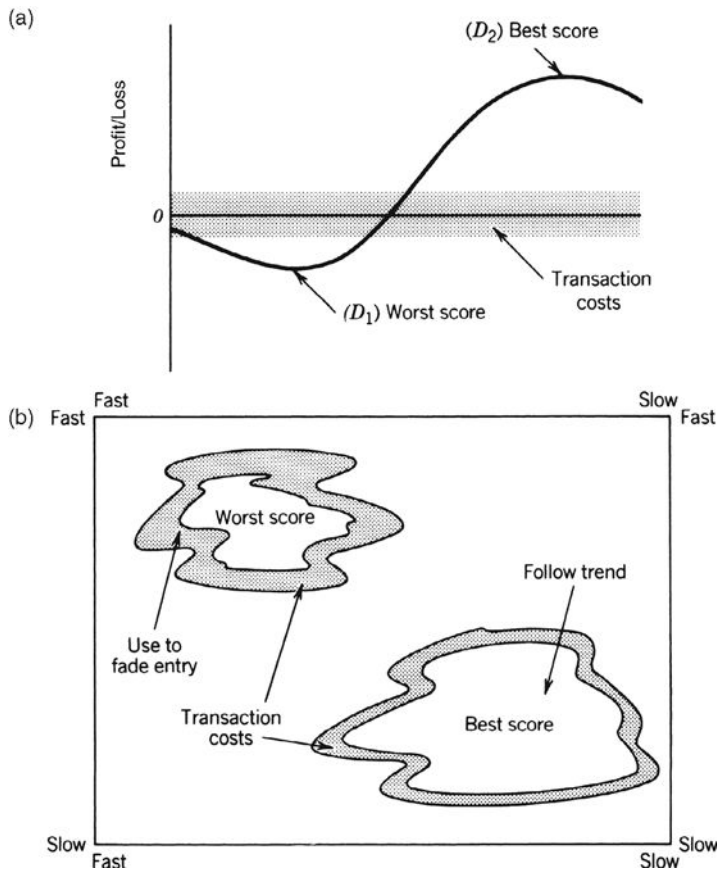
FIGURE 21.16   Dual use of test map. (a) Single-variable test.
(b) Two-variable test.

Following this strategy, performance is greatly improved over the simple trend. Although the *Net PL* is lower, all the other statistics are far better, as seen in the table below. In addition, trades are not reversed from long to short, so that liquidity is improved and some unexpected risk may be avoided.

Strategy	Net PL	Profit Factor	Trades	%Prof Trades	Profit/Trade
MA	64385	1.51	212	35	277
MA w/Fade	53380	2.26	72	47	741

RETESTING FOR CHANGING PARAMETERS

One optimization is never the end of the testing process. If the system development is completed, the parameter values selected, and the most recent data used for out-of-sample verification, then you have *proved the concept*. One final test must be done on all of the data before beginning to trade the program. Using more data always gives a more

robust solution. The final parameter choices may be slightly different from the previous values but should not change significantly. A large change is a warning of problems, either a test period that was too short, or a poor selection of parameter values.

As in step-forward testing, some analysts prefer to retest a system periodically using a fixed amount of data, dropping off the oldest. Longer test intervals are always better, unless there is a reason to believe that the oldest data is no longer relevant. In some cases, there are structural changes, such as the formation period preceding the emergence of the European Currency Unit, which put controls on the variance in exchange rates between the participating countries. A new market with low liquidity during its earlier years may be considered different from more active, current market conditions. Retesting remains an important way to adjust the system to characteristics of new data.

Retesting should be considered after 5% to 10% of new data is available, relative to the original test interval. Using the 2-variable test as an example, a shift may be expected in the area of best performance as shown in Figure 21.17.

Because only a small amount of data was added, the size of the shift in Figure 21.17 should be small. If a large shift actually occurs, then two possibilities should be considered:

1. The data was unusually volatile and introduced patterns not previously seen in the data. In this case the shift in parameters is justified.
2. The data period for the original tests was short and did not include enough patterns to make the parameter choice robust. With a small amount of test data, there is a
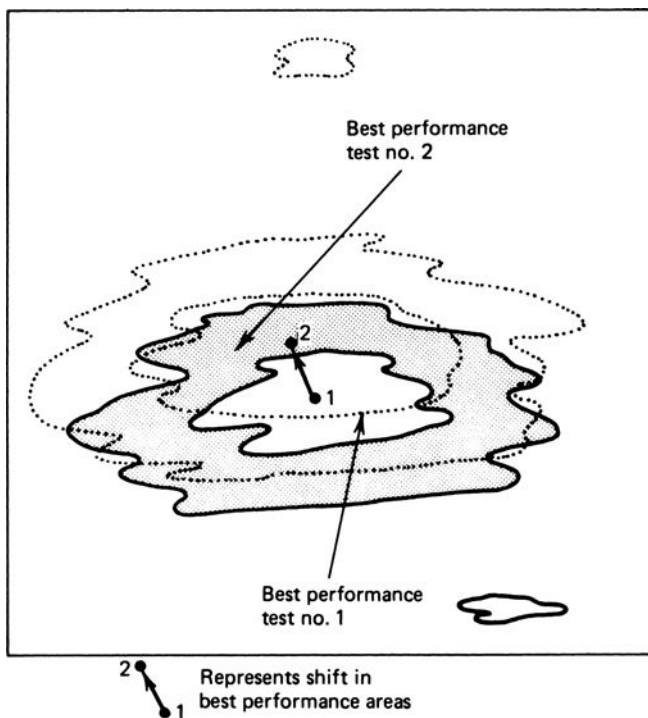
FIGURE 21.17   Consecutive tests.

greater likelihood that a faster trading model will be selected. As the data shifts and that faster model is no longer good, the parameters may jump to a very slow model.

Retesting after a Major Move

When your system fails to perform as expected, often after a price shock, it is natural to question whether the system has the right parameters. Is the trend fast enough? Is the stop-loss too close? If you have used enough data for your original tests, then it is very likely that retesting won't change anything. If you make up new rules to fix this one problem, you are overfitting and therefore accomplish nothing. While using more data for testing is always safest, markets will continue to offer new patterns, bigger swings, and more risk than anything contained in the history.

Is the Model Broken?

Some trading systems, when based on an economic scenario, will have a limited life. A fully technical program, such as a moving average, should adapt to trends of varying lengths that occur over time, taking larger and smaller profits, and cutting losses short. The way to judge if a model is performing as expected is by measuring the size and duration of drawdowns and comparing that to the tested profile, understanding that, as the amount of data increases, so will the likelihood of larger profits and larger losses. If a new system begins trading and posts steady losses until it reaches the previous maximum drawdown, chances are that something is wrong. Systems should not systematically decay from the first day of trading.

The probability of a drawdown greater than P can be calculated as[11]

$$ Probability \ of \ a \ drawdown \geq P = \left(1 - \frac{P}{100}\right)^{2\mu/\sigma^2} $$

This formula assumes a normal distribution of returns (although trading returns should have a longer tail to the right). The value μ is the average fractional change in equity (percentage returns), and σ is the standard deviation of the returns. The equation gives the probability that the daily drawdown from the peak equity will be greater than or equal to P. The same formula works for monthly data. For example, during the test, the annualized return was 30% with a daily return ratio of $\sigma/\mu = 9$. Based on 250 trading days per year, the daily value of μ is calculated using $(R_y/R_0)^{1/y} - 1$, or $\mu = (1.3/1.0)^{1/250} - 1$, and $\mu = 0.00105$. Then $\sigma = 0.00945$ and the value of the exponent $2\mu/\sigma^2 = 23.51$. If we want the 95% confidence level, then the final step is $P = (1 - 95/100)^{1/23.51} = 0.880$. Then the probability of a drawdown of at least 12% can be expected in 5% of the months.

Chapter 23 has other measures of risk, including the use of residuals to find the probability of a drawdown over a single time period, and the probability that equity will remain below a threshold level for a specified time interval.

---

[11]John R. Wolberg, *Expert Trading Systems* (New York: John Wiley & Sons, 2000), 181–183.

TESTING ACROSS A WIDE RANGE OF MARKETS

A comprehensive test that applies one trading strategy to a broad selection of markets is basic to the concept of *robustness*. When a system performs well in many markets over the same test period and calculation ranges, it is fair to assume that the method is sound. If a technique works in the Swiss franc but not the euro, in 10-year Treasury notes but not 5-year Notes, in Intel but not Microsoft, then the method is most likely not robust but fine-tuned to each market. Markets have individual characteristics that might justify differences in performance; however, a strategy that can generate profits in a broad range of markets with the same rules and calculations is robust. At the same time, we should expect the results to be *suboptimal*; that is, performance will vary considerably across markets because fewer parameters and rules will be used. As Ken Tropin, founder of Graham Capital says, "Loose pants fit everyone." In addition to the studies that follow, Chapter 8 includes a section, "Comparison of Major Trend Systems," that will be helpful in viewing the results of different trend-following techniques side by side.

Robert Colby, in *The Encyclopedia of Technical Market Indicators*,[12] has given an excellent comparison of market indicators by presenting each of their test results in a common form. Thomas Bulkowski has done the same in his *Encyclopedia of Chart Patterns*.[13] Viewing the results of a set of indicators gives a much better idea of the success of that approach than testing any one of them. For example, if one volume indicator was successful during the past 10 years, but three others were inconsistent, then the use of any volume indicator becomes questionable. While there are always exceptions, robustness is found when a number of strategies, all applied to volume, show similar, profitable returns. This was the conclusion when a selection of trending methods was tested in Chapter 8.

Classic Tests by Davis and Thiel, Maxwell, and Hochheimer

There are three previous (if somewhat dated), well-known comprehensive studies of moving average systems by Davis and Thiel, Maxwell, and Hochheimer. In addition, Lukac, Brorsen, and Irwin studied 12 different trading methods across 12 futures markets. These studies are interesting because of the techniques and the conclusions. Strategy development software now allows us to test, in seconds, what once took hours or days, so the actual results of these older studies will be omitted in favor of our own tests.

Davis and Thiel

Davis and Thiel, the authors of the oldest of the studies,[14] completed before 1970, analyze the greatest variety of markets, covering virtually all of the U.S. crops as well as cattle, eggs, and the soybean complex. They include about 5 years of data and use relatively

---

[12] Robert W. Colby, *The Encyclopedia of Technical Market Indicators* (New York: McGraw-Hill, 2003).

[13] Thomas Bulkowski, *Encyclopedia of Chart Patterns* (New York: John Wiley & Sons, 2000).

[14] R. E. Davis and C. C. Thiel, Jr., *A Computer Analysis of the Moving Average Applied to Commodity Futures Trading* (West Lafayette, IN: Quiatenon Management, 1970).

fast moving averages (up to 10 days); they introduce variations in lead-oriented plotting and in testing nonconsecutive days; the data used is close-only. The results are clearly presented in both detail and summary, and yield generally good returns.

The study used combinations of simple moving average buy and sell signals, leading plots (moving the calculation forward), and skips (using every second or third price rather than every price). They test a total of 100 combinations of the three factors:

A *skip* from 1 day (none) to 5 days (1 week).
A *moving average* of 5, 6, 7, 8, or 10 days.
A *leap* of 0 or 2 days.

Before 1970, markets had fewer participants, less noise, and showed more short-term trending, and the moving average was state-of-the-art technology. Using a computer to study the markets in 1970 was very unusual; Digital Equipment was still selling their most popular PDP-5, and IBM had just introduced its first mini-computer, the System 3. Traders were using point-and-figure charts on the exchange floors, and the 10-day moving average was the most popular technical tool. Lower volume and a larger percentage of commercial participation made trends more consistent than today's markets. It was possible to trade using a 3- to 5-day moving average and post consistent profits.

In retrospect, the technique that skipped 1 or 2 days of data is another way of smoothing, similar to a 3-day momentum that compares today's price with the price 3 days ago. The success of the variations in rules is proof that markets trended in the 1960s and that a trending method, such as a moving average, was a viable strategy. Over time it has been necessary to increase the calculation period in order to avoid steadily increasing noise.

Maxwell

Maxwell's study is extremely comprehensive but applied only to the pork bellies futures market.[15] His idea was to perform a thorough test of combinations of trending methods and trading rules. As with Davis and Thiel, the computer was still not available to the general public, and the ability to perform a broad study was a groundbreaking event.

Trend combinations included a simple moving average of 3, 5, or 10 days, an average-modified or a weighted average. The second feature was the delay factor, used to improve the timing of both entries and exits. Some of the options were:

1. Act without delay.
2. Act if the signal condition persists for one additional day (a confirmation).
3. Enter if the signal condition persists for two additional days, but liquidate without a delay.

Combinations of two types of moving averages and a delay factor were tested, with and without fixed or moving stops. With 10 types of averages, 6 delay factors, and different stops, Maxwell had a lot of combinations to examine.

---

[15]Joseph R. Maxwell, *Commodity Futures Trading with Moving Averages* (Cupertino, CA: Speer, 1974).

The study is then expanded to 3-factor systems with a list of 18 combinations of rules to generate 324 systems of which the results of 285 are recorded, with 49% profitable systems. The largest *loss* was generated by a system with the rules:

1. Enter a new position when the price crosses both the weighted 3-day and weighted 10-day averages, as long as the signal condition persists for two additional days (*buy* if prices are higher, *sell short* if lower).
2. Liquidate positions when the 3-day average reverses its direction through the longer average as long as it continues for one additional day.
3. No fixed stops are used.

The best *profits* in the 3-factor system used the rules:

1. Enter a new position when both the 5-day average-modified and a 10-day weighted average move above the average price, provided the 5-day average lagged behind the 10-day (that is, the longer-term trend turned up first in order to get a buy signal).
2. Liquidate the position when the shorter-term average crossed the longer-term average in the direction opposite to the current position.
3. No fixed stops were used.

Maxwell's study represents a great amount of work and some simple and sound philosophy. The conclusion—that two trends, one twice the length of the other, will be profitable if the slow trend turns first—makes sense even with longer calculation periods in current markets. The slower trend finds the major move and the faster trend provides the timing. Identifying the major move comes first. In the rules that produced the worst results, the 2-day lag to confirm the signal is long with respect to the 3- and 10-day trends. When trading using fast signals, a delay will often reduce the net profits.

Unfortunately, Maxwell's studies only used a 50-day test interval for May 1972 pork bellies, which is not an adequate data sample to statistically justify his conclusions. Maxwell does test four other selected 50-day periods, but the reader cannot know how these periods were chosen. Considering the effort in outlining a testing program and establishing rules, in addition to the analysis of results, the test period is a very weak point. Because the final test rules and values are reasonable, the market may have been successfully traded with a broad range of trending strategies.

Comparing Methods of Calculation across Selected Markets

Hochheimer[16] performs two interesting studies: a comparison of three types of *moving average calculations* and a test of *channels* and *crossovers*. In the first analysis, Hochheimer compares *simple*, *exponentially smoothed*, and *linearly weighted* moving

---

[16]Frank Hochheimer, "Moving Averages," and "Channels and Crossovers" in Perry J. Kaufman (Ed.), *Technical Analysis in Commodities* (New York: John Wiley & Sons, 1980).

averages, tested on a good sampling of markets (without financials and currencies, which did not exist during the 1970–1976 test period). The calculations for simple moving average and exponential smoothing can be found in Chapter 8. The *linearly weighted average* is actually *step-weighting*, where integer values are used as weighting factors, each successive day is incremented by one, and the most recent day has the highest weighting factor.

The same trading rules applied to all trends, regardless of the calculation method:

1. The trend calculation used the closing price only.
2. A *buy* signal occurred on an upwards penetration of the moving average by the closing price; a *sell* signal occurred on a downward penetration of the moving average.
3. The model is always in the market.
4. A trade cannot be executed when the day's high and low price are the same. This assumes that there was no trading, caused by a *locked-limit day*, the consequence of a price shock.

The test covered 7 years of data from 1970 to 1976 and calculation periods which ranged from 2- to 70-day periods (or the equivalent smoothing constant). In Hocheimer's results, the slower trends were consistently better than faster ones, regardless of which technique was used. With a few exceptions, the best selection of days ranged from 40 to 70. This may be the first sign of a shift to longer calculation periods, in contrast to the approach taken by Davis and Thiel as well as Maxwell.

Current Results

To bring this up to date, a sample of 17 liquid futures markets were tested for five different trend-following strategies using calculation periods from 2 to 80 days, in steps of 2 days, and data from 1990 through mid-2011. The methods are

1. Moving average
2. Exponential smoothing
3. Linearly weighted average
4. Slope of a least-squares regression
5. *N*-Day breakout

All trades were charged a round-turn cost of $40 to cover commissions and slippage. Where markets were not yet trading at the beginning of the test period, data was used from the inception of that market. Table 21.3 gives a summary of the moving average results. The columns that are not obvious are:

> *Best Period*, the calculation period with the highest Net Profit of all tests.
> *Best Range*, the range calculation periods showing the best test performance, based on observation.

TABLE 21.3 Results of a Simple Moving Average Model

Market	Best Period	Best Range	Net PL	Profit Factor	Trades	% Prof Tests
U.S. bonds	72	62–80	36291	1.25	228	30%
Eurobund	80	48–80	55540	1.87	156	73%
Eurodollars	28	22–80	33867	1.64	289	78%
S&P	72	64–72	3197	1.02	288	5%
DAX	42	58–72	213135	1.51	318	80%
NASDAQ	70	44–78	80785	1.87	162	73%
EURUSD	74	24–36	169332	1.80	216	85%
USDJPY	66	62–80	98295	1.50	256	75%
CADUSD	64	38–68	9670	1.08	290	3%
Crude oil	20	64–80	161070	1.50	152	85%
Heating oil	74	16–80	174687	1.85	247	75%
Natural gas	24	58–80	692640	1.65	132	90%
Gold	50	14–24	11830	1.07	322	3%
Copper	74	40–50	153137	2.03	250	80%
Wheat	14	60–80	23420	1.12	556	10%
Cotton	80	10–70	106855	1.97	223	83%
Soybeans	24	58–80	33970	1.15	466	10%
Average	**55**	**22–26**	**121042**	**1.52**	**268**	**55%**

Profit Factor, the gross profits divided by the gross losses. This measure is the only risk measure available on TradeStation and does not always show the relative return and risk correctly. It is based on closed-out trade results; therefore, any large losses held during the trade are ignored if that trade closes out at a better level.

% Prof Tests, the percentage of profitable tests out of the total of 40 calculation periods. This gives us a notion of the robustness of the method. If 30 of the 40 tests were profitable, regardless of the amount, the result is 75%.

Caveats

The test range selected was 2 to 80 days, which can be criticized as an unrealistic choice. First, we do not expect profits using a 2-day trend because it is too fast, and generates too many trades with small profits and losses that will not be greater than the cost of $40. Nevertheless, all of the systems that will be compared suffer the same burden. Second, the number of days in the test should not be evenly distributed because the difference in performance between a 2- and 4-day trend is far greater than the difference between a 78- and 80-day trend. However, it is more complex to create tests this way, and the intention here is to present something that can easily be replicated by any analyst with modest system development software.

Moving Average Conclusions

We can draw some useful conclusions from this first test.

1. The best choice is often near the high (longer calculation periods) end of the test range.

2. The best range of results often includes the maximum test period, 80 days.

3. Those markets that showed unusual results can, for the most part, be explained.

4. The percentage of profitable tests, a measure of robustness, is high for most markets.

This confirms that a moving average is a viable strategy and the slower calculation periods tend to be uniformly better.

Of all test results, robustness *(% Prof Tests)* is the first one to be reviewed. A market that does not trend has few tests that are profitable. The worst of those markets are the S&P, the Canadian dollar, and gold. Of these, gold may be the most surprising because it has increased dramatically during the past few years; however, it languished at low prices for 15 of the 20 years tested. If a closer look at the results of gold showed that it was highly profitable for the past 5 years, you might want to include it based on the assumption that volatility will continue and that it adds important diversification to a portfolio.

The Canadian dollar does not have the impact of gold. It reflects a decline in the value of the U.S. dollar, but is often subdued in price movement and, in many cases, has a difficult time overcoming costs. It also has made bigger moves recently, but may not offer enough positive arguments to offset its poor historic performance.

The S&P presents the most difficult problem. It is a very important, liquid market, but has been shown to have a large component of noise that fights with the trend. Robustness of 5% means that only 2 of 40 tests were profitable, and those do not necessarily show profits recently. While the DAX and NASDAQ were both very successful under this strategy, it is difficult to argue why the S&P should be included. Each analyst will need to resolve this dilemma himself.

One way to get to see the big picture of the results is to chart the average net profit and average profit factor for all 17 markets by their calculation periods. Figure 21.18 shows that the faster trends are uniformly losses, while progressively longer trends are profitable. This argues for the success of the *macrotrend*, which tends to track the bigger price moves caused by economic policy. It also shows that the profit factor does not offer much additional information on risk than the net profits.

Exponential Smoothing

Smoothing is different from a moving average because it is front-loaded; that is, it puts more weight on the nearest data while decreasing the importance of older data each day. This is discussed in Chapter 7. One important characteristic is that the trend always turns up when prices move above the trendline and turns down when they move below
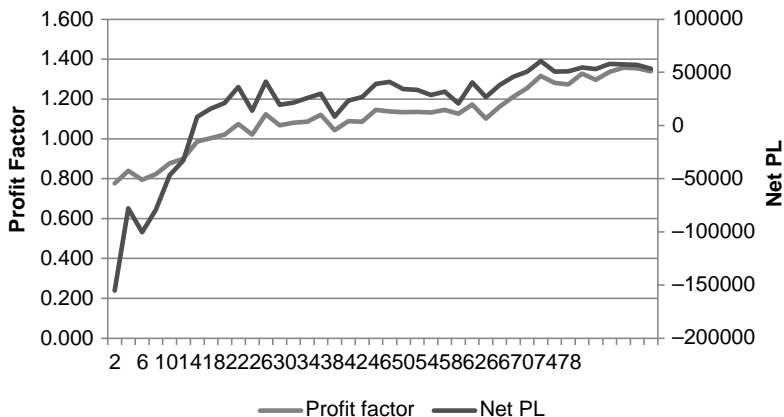
FIGURE 21.18   Average results of moving average tests for 17 markets based on calculation period, showing the similarity between the profit factor and the net PL.

the trendline. This increases the number of trades and offers no advantage to using the direction of the trendline as the buy and sell signal rather than the price penetration of the trendline. Table 21.4 summarizes the results in the same form as the moving average tests. Other than having more trades, these results are very similar to the moving average.

TABLE 21.4   Results of Exponential Smoothing Tests

Market	Best Period	Best Range	Net PL	Profit Factor	Trades	% Prof Tests
U.S. bonds	78	70–80	43068	1.32	230	35%
Eurobund	78	64–80	73140	2.64	134	70%
Eurodollars	80	24–80	28310	1.72	178	78%
S&P	60	58–74	555	1.00	319	3%
DAX	12	12–22	228282	1.31	651	90%
NASDAQ	70	58–68	67260	1.71	160	75%
EURUSD	68	38–80	206447	2.08	223	83%
USDJPY	54	60–76	125892	1.63	294	83%
CADUSD	66	38–64	12810	1.11	314	10%
Crude oil	42	62–66	140650	1.56	329	88%
Heating oil	72	38–80	201814	2.07	251	78%
Natural gas	28	48–80	266180	1.62	369	93%
Gold	70	26–46	15210	1.09	306	13%
Copper	74	68–76	154582	2.13	216	85%
Wheat	58	56–80	1840	1.01	297	3%
Cotton	80	56–62	135495	2.51	223	75%
Soybeans	20	56–80	15175	1.05	580	8%
Average	**59**	**18–22**	**100983**	**1.62**	**298**	**57%**

TABLE 21.5 Results of Linearly Weighted Average Tests

Market	Best Period	Best Range	Net PL	Profit Factor	Trades	% Prof Tests
U.S. bonds	18	10–20	55215	1.16	627	30%
Eurobund	80	56–80	42140	1.46	248	63%
Eurodollars	40	30–52	22957	1.38	311	73%
S&P	62	62–74	50662	1.24	405	20%
DAX	62	60–72	267080	1.63	319	78%
NASDAQ	68	50–80	72605	1.56	232	63%
EURUSD	30	26–32	142825	1.31	475	83%
USDJPY	44	64–80	98847	1.31	418	73%
CADUSD	64	40–66	35690	1.29	385	30%
Crude oil	80	62–68	157190	1.75	281	80%
Heating oil	74	64–80	132282	1.52	319	80%
Natural gas	18	44–78	244480	1.34	634	68%
Gold	78	10–20	41300	1.27	318	13%
Copper	70	78–80	178655	1.93	329	88%
Wheat	10	42–80	−15525	0.94	850	0%
Cotton	80	68–80	118475	1.84	301	83%
Soybeans	34	60–80	9205	1.03	469	5%
Average	**54**	**34–56**	**97299**	**1.41**	**407**	**54%**

Linearly Weighted Average

The linearly weighted average is also frontloaded, giving the largest weight to the most recent data. Unlike exponential smoothing, the oldest data does not retain any residual value, but drops off at the end of the calculation period. Table 21.5 shows the test.

Linear Regression Slope

The linear regression slope is popular among economists, but the method is very different from the previous three averages. The regression line slope shows the trend of prices that continue in the same direction and will trigger a signal sooner than a moving average. This can be seen in the results of the Canadian dollar and the S&P, although wheat and soybeans do not perform well. (See Table 21.6.)

N-Day Breakout

This last trending method is very different from the others because it does not have a "trendline" but identifies the trend based on new highs and lows. The *N*-day breakout is deceiving in that it is both very reliable in finding the correct trend but at exceptionally high risk. The risk cannot be seen in these numbers; however, every trade has an initial risk equal to the highest high minus the lowest low over the calculation period. For the longest period of 80 days, that risk can be very large and is the reason why many traders prefer the basic moving average over the breakout. Table 21.7 shows the results.

TABLE 21.6 Results of Linear Regression Slope Tests

Market	Best Period	Best Range	Net PL	Profit Factor	Trades	% Prof Tests
U.S. bonds	80	16–32	68377	1.99	57	65%
Eurobund	80	60–80	54420	2.52	46	83%
Eurodollars	50	28–60	45632	2.73	76	85%
S&P	74	74–08	20550	1.19	75	28%
DAX	74	66–76	239747	2.19	63	88%
NASDAQ	80	64–80	87655	2.94	38	80%
EURUSD	38	28–44	173272	1.99	54	93%
USDJPY	76	64–80	119555	2.17	59	88%
CADUSD	72	64–80	44050	1.74	75	33%
Crude oil	26	20–34	146150	1.67	173	83%
Heating oil	70	66–80	176939	2.36	74	70%
Natural gas	24	16–36	250020	1.64	201	88%
Gold	26	22–40	−1550	0.99	177	0%
Copper	80	56–80	151302	2.68	57	83%
Wheat	24	22–34	18557	1.14	206	18%
Cotton	80	60–80	77110	1.79	65	88%
Soybeans	26	22–44	44022	1.27	208	38%
Average	**58**		**100930**	**1.94**	**100**	**65%**

TABLE 21.7 Results of *N*-Day Breakout Tests

Market	Best Period	Best Range	Net PL	Profit Factor	Trades	% Prof Tests
U.S. bonds	62	60–80	121698	1.85	41	70%
Eurobund	62	58–74	83290	6.86	29	90%
Eurodollars	80	56–80	52175	5.10	25	46%
S&P	58	46–60	16007	1.17	51	38%
DAX	64	62–80	260477	2.69	37	75%
NASDAQ	40	16–50	149510	2.66	46	85%
EURUSD	32	22–40	157917	1.96	79	90%
USDJPY	60	56–70	100355	2.08	44	88%
CADUSD	68	62–80	48630	2.75	32	53%
Crude oil	60	56–80	134920	2.09	46	88%
Heating oil	60	50–80	174513	2.48	44	85%
Natural gas	80	14–24, 68–80	247060	3.08	33	39%
Gold	18	14–18, 74–80	11720	1.08	161	13%
Copper	44	26–56	127400	2.22	60	78%
Wheat	80	66–80	37530	2.06	34	38%
Cotton	78	64–80	91645	2.65	31	93%
Soybeans	24	22–28	26645	1.20	126	35%
Average	**57**		**108323**	**2.59**	**54**	**65%**

Comparing the Five Strategies

We can draw some important conclusions from looking at the average results of the five trend-following tests, shown in Table 21.8. The first and most significant point is that the averages are all profitable. That means trend-following works. If prices trend, then any strategy that profits from the trend should work to some degree. It is not the method, but the market that is most important. Each strategy extracts slightly different parts of the same trends.

Of the five methods, the weakest is the linearly weighted average. It has the most trades and the lowest robustness, although only 1 point below the moving average. The two that show the best performance are the linear regression slope and the breakout, but they also have far fewer trades and much higher risk. An average of 54 trades over 20 years means that the breakout system had only 2.7 trades per year. In some years it may have had 5, and in others it held the same trade all year. That translates into high risk. Then the breakout system should reward you for taking that risk, yet the basic moving average system had the highest average net profit without that exposure.

The next big picture item is robustness. If a market trends, then we can expect any of the trend systems to be profitable for that market. Certain characteristics of a strategy will cause the results to vary. For example, the two extremes are the simple moving average, which cuts losses short, and the breakout, which holds the trade as long as possible. For the S&P, the moving average posted only 2 tests of 40 with profits, while the breakout had 15 profitable tests. By looking at the average of all systems, the far right column in Table 21.9, we can see which markets are trending and are likely to perform best with a trend-following strategy.

Sorting the results in the right column of Table 21.9 and charting them in Figure 21.19, we can see that there is a sharp delineation between trending and non-trending markets. Of the 17 markets, 11 show robustness above 70%, which would be a good threshold for acceptance. However, a decision cannot always be made on a single criterion. For example, interest rates, represented by U.S. 30-year bonds, reflect interest rate policy, and have been the biggest component of profits for macrotrend portfolios for the past 20 years. They consistently produce profits using long-term trends but fail for faster calculation periods. As mentioned earlier, gold has been very profitable during the past few years and has increased in volatility. If gold continues to be an alternate currency choice for investors, then performance during the 1990s may not be important.

TABLE 21.8   Average of Statistics of Five Trend-Following Strategies

Strategy	Best Period	Net PL	Profit Factor	Trades	% Prof Tests
MA	55	121042	1.52	268	55%
EXP	59	100983	1.62	298	57%
LWA	54	97299	1.41	407	54%
LRS	58	100930	1.94	100	65%
BO	57	108323	2.59	54	65%

TABLE 21.9   Average Robustness of Trend Tests by Market

Market	MA	EXP	LWA	LRS	BO	Average
U.S. bonds	30%	35%	30%	65%	70%	46%
Eurobund	73%	70%	63%	83%	90%	76%
Eurodollars	78%	78%	73%	85%	46%	72%
S&P	5%	3%	20%	28%	38%	19%
DAX	80%	90%	78%	88%	75%	82%
NASDAQ	73%	75%	63%	80%	85%	75%
EURUSD	85%	83%	83%	93%	90%	87%
USDJPY	75%	83%	73%	88%	88%	81%
CADUSD	3%	10%	30%	33%	53%	26%
Crude oil	85%	88%	80%	83%	88%	85%
Heating oil	75%	78%	80%	70%	85%	78%
Natural gas	90%	93%	68%	88%	39%	75%
Gold	3%	13%	13%	0%	13%	8%
Copper	80%	85%	88%	83%	78%	83%
Wheat	10%	3%	0%	18%	38%	14%
Cotton	83%	75%	83%	88%	93%	84%
Soybeans	10%	8%	5%	38%	35%	19%
Average	**55%**	**57%**	**54%**	**65%**	**65%**	**59%**

Of the remaining markets, wheat and soybeans have not been discussed. In the same way that interest rates perform only for longer-term trends, grains are dominated by seasonality. The very long-term trend will identify structural changes as well as shifts in the U.S. dollar, but profits from that are minimal compared to seasonal price swings.

Crossover Strategy

In Hochheimer's *crossover of two simple moving averages*, the short-term average (the faster trend) ranged from 3 to 25 days and the longer one from 5 to 50 days. The objective
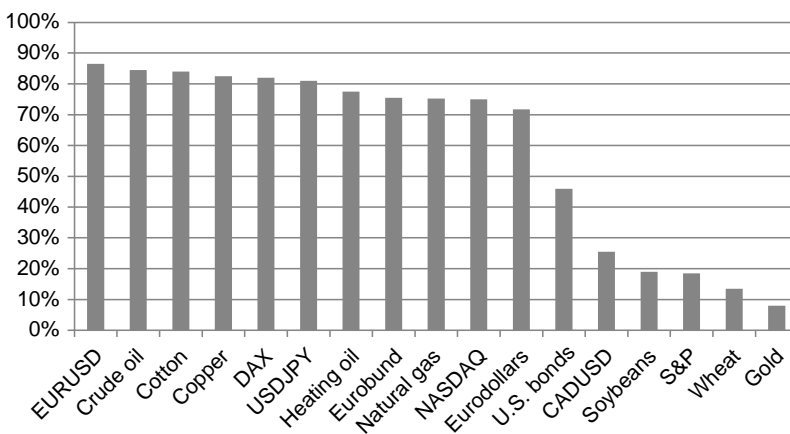
FIGURE 21.19   Robustness of trend-following tests, by market.

was to eliminate the whipsaws that were evident in the first study. The rules for this system were:

1. Each moving average used closing prices only; the calculation period for the long-term trend was always greater than the period for the short-term trend.

2. *A buy* signal occurred when the short-term average moved above the long-term average; a *sell short* signal occurred when the short-term average moved below the long term.

3. The model was always in the market.

4. If the high and low prices of the day were equal, a *locked-limit* day was assumed, and no trading signals were generated.

In the current test of two trends, the short period will vary from 3 to 21 in steps of 2, and the long period from 10 to 80 in steps of 5. That gives 150 combinations of which 11 are eliminated because the short period will be greater than the long period (for example, the combination of a 15-day fast trend with a 10-day slow trend is not allowed).

Results are more difficult to display because each test is a 3-dimensional surface, as shown in Figure 21.20 for the DAX index. In general, the best results favor the longest calculation periods of both parameters, seen as the rounded area running up the front right of the chart. The conclusion here is that this method does best with the slowest signals, the same conclusion as the single trend approach.

The results of testing the same set of markets are shown in Table 21.10. Notice that the average profit factor is lower than any of the single-trend strategies and the average number of trades is higher than all but the linearly weighted average. A high number of trades is likely to mean smaller profits per trade. Traders have a minimum unit return below which the return is unrealistic. In Table 21.11, the per contract returns show that
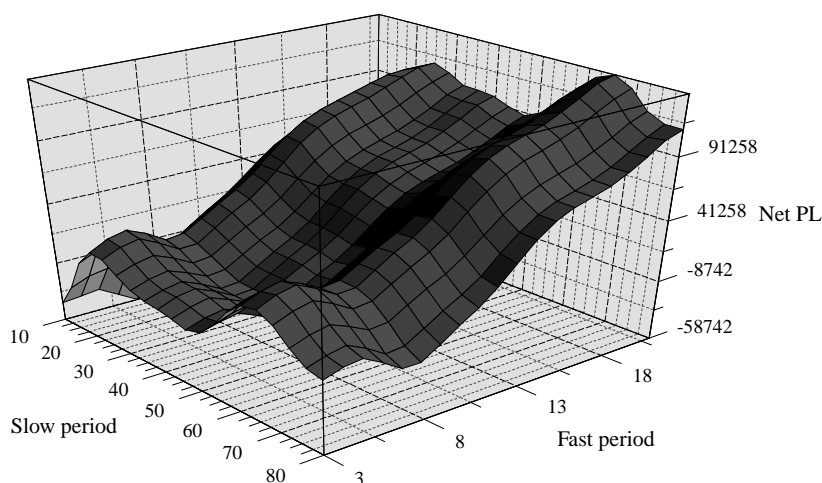
FIGURE 21.20 Surface plot of the DAX index optimization results. (Plot created with PSI-Plot by Poly Software International)

TABLE 21.10 Results of a Moving Average Crossover Strategy

Market	Best Slow	Best Fast	Net PL	Profit Factor	Trades	% Prof Tests
U.S. bonds	80	17	35265	1.19	342	27%
Eurobund	75	17	53270	1.60	277	55%
Eurodollars	30	21	23685	1.48	299	56%
S&P	75	13	2290	1.01	452	1%
DAX	60	21	185795	1.48	326	73%
NASDAQ	70	17	53415	1.41	238	41%
EURUSD	25	19	131607	1.36	426	89%
USDJPY	65	15	59682	1.24	401	54%
CADUSD	75	15	–32020	0.77	394	0%
Crude oil	65	19	156120	1.75	328	63%
Heating oil	70	19	158159	1.58	356	53%
Natural gas	20	17	27396	1.58	423	91%
Gold	50	19	22600	1.13	381	6%
Copper	75	19	114415	1.56	352	72%
Wheat	15	13	11452	1.06	527	9%
Cotton	80	15	74180	1.48	381	84%
Soybeans	25	21	26432	1.12	476	8%
Average	56	18	64926	1.34	375	46%

TABLE 21.11 Per Contract Returns for the 2-Moving Average, the Simple Moving Average, and the Linear Regression Slope Strategies

Market	2 MA	MA	LRS
U.S. bonds	103	159	1199
Eurobund	192	356	1183
Eurodollars	81	117	600
S&P	5	11	274
DAX	569	670	3805
NASDAQ	224	498	2306
EURUSD	308	783	1533
USDJPY	148	383	2026
CADUSD	–81	33	587
Crude oil	475	354	844
Heating oil	444	707	2391
Natural gas	647	753	1243
Gold	59	36	–8
Copper	325	612	2654
Wheat	21	42	90
Cotton	194	479	1186
Soybeans	55	72	211
Average	222	357	1301

the average for the 2-trend system is only 62% of the simple moving average and 17% of the linear regression slope. And, while $222 is large enough overall, the Eurodollar interest rates show only $81, which is at the low end of what most analysts would regard as a safe margin of error.

The measure of robustness is also lower than any other system, indicating that nearly half of all moving average pairs that were tested generated losses. For the S&P and Canadian dollar, there were essentially no combinations that were profitable, but then the ranking of successful markets is similar to the previous tests. In reviewing the single-trend systems, we could conclude that, in general, they favored the longer calculation periods. In the 2-trend system, there is also a tendency to pick the longest of the fast trend, which had a maximum calculation period of 21 days. Had we allowed both trends to assume the same range up to 80 days, it is very likely that some of the combinations would have chosen both trends at 80, or very close to it. This method has forced us to choose a suboptimal combination because, from the robustness measurement, the 2-trend method is not as good as any of the single-trend strategies. To verify that observation, we would need to measure the information ratio for all of the systems, keeping in mind that the average results should be the criterion for comparison, and not the best results.

More on Comparing Strategies and Markets

The key to robustness is to test more data, more markets, more of everything, and then to study the results looking for a common areas of success and broad patterns of consistency. There is no better way to tell if a market has trends than to test it for a variety of trending methods. A market that performs well using a daily breakout, but fails with a regression slope, is not a good candidate for trend system.

A comprehensive look at 12 trading methods applied to 12 futures markets provides an opportunity to evaluation generalized performance.[17] The annualized percentage returns for a variety of U.S. markets are shown for 12 well-defined technical methods in Table 21.12. These trading systems are:

CHL *Channel breakout*, the closing price penetration of the *N*-day high and low.

PAR Wilder's *Parabolic system*, in which the stop and reverse point (SAR) gets closer each day.

DRM *Directional Movement*, a Wilder method that averages the up days and down days separately.

RNQ *Range Quotient*, basing a breakout of a ratio of current change to past change.

DRP *Directional Parabolic*, combining the DRM and PAR systems.

MII *MII Price Channel*, a breakout system using only the oldest and most recent prices.

LSO *L-S-O Price Channel*, using two parameters, including an interval of oldest prices, to decide a channel breakout.

---

[17]Louis P. Lukac, B. Wade Brorsen, and Scott H. Irwin, "How to Test Profitability of Technical Trading Systems," *Futures* (October 1987).

TABLE 21.12 Percentage Returns by System and Market, 1978–1984*

	CHL	PAR	DRM	RNQ	DRP	MII	LSO	REF	DMC	DRI	MAB	ALX	Average
Corn	22.4	3.8	29.8	−6.2	30.1	43.3	10.7	4.6	37.9	−2.9	−4.7	22.9	16.0
Cocoa	10.0	−101.7	−121.9	−345.0	−112.1	−73.6	−256.9	−120.5	−72.9	−281.8	−219.2	−35.2	−144.2
Copper	−15.4	2.9	−46.1	−78.4	−4.2	−31.4	−83.0	−66.2	−39.8	−94.3	−118.4	−16.8	−49.3
Cattle	−12.2	−28.4	−12.4	−72.5	−16.5	−34.8	−44.9	−56.3	−11.4	−70.0	−58.9	−5.5	−35.3
Pork bellies	−30.8	22.0	−6.8	−134.4	4.8	20.6	−117.4	−145.3	−6.5	−127.6	−112.0	−12.1	−53.8
Lumber	38.7	−43.6	−0.4	−19.2	−40.6	31.3	−46.3	−18.4	36.6	−19.1	−47.1	24.6	−8.6
Soybeans	7.7	−19.1	25.8	−57.9	−10.0	13.3	−21.5	−45.6	2.6	−10.7	−22.5	13.1	−10.4
Silver	60.5	54.4	0.2	−15.9	82.3	−19.1	12.5	−72.4	34.2	−76.2	−18.0	55.9	8.2
Sugar	103.4	46.2	61.2	−42.1	63.4	72.6	−0.8	−47.3	82.3	17.6	15.2	71.6	36.9
British pound	−3.5	8.1	20.7	−36.1	30.3	1.9	−7.8	3.5	25.9	−10.2	−65.8	−39.8	−6.1
German mark	66.5	35.4	68.2	18.8	78.0	63.3	19.2	−17.8	46.3	24.6	6.6	−50.0	29.9
T-bills	108.7	48.5	132.7	−40.1	225.5	189.9	221.8	239.8	127.8	39.2	−12.0	32.9	109.6
Average	29.7	2.4	12.6	−69.1	27.6	23.1	−26.2	−28.5	21.9	−51.0	−54.7	5.1	−8.9

*Calculation of annual percent returns assumes that *30%* of the initial investment is used for initial margin, in which margin is estimated as 10% of contract value.

REF *Reference Deviation*, determines a volatility breakout based on a standard deviation of past closing price changes.

DMC *Dual Moving Average Crossover*, holding a trend position when both moving averages are moving in the same direction.

DRI *Directional Indicator*, a ratio of the current price change to the total past price change.

MAB *Moving Average with % Price Band*, gives a signal when prices penetrate the band.

ALX *Alexander's Filter Rule*, generates signals when prices reverse from a previous swing high or low by a percentage amount.

To create the results in Table 21.12, each market was tested for 3 years and the best parameters used to generate the returns for the next years, in a step-forward approach. It is easy to see that some markets are not profitable for any strategy and that some strategies are generally poor. Unless you were creating a strategy for a specific group of markets, you would not want to trade a system that was not profitable in at least 50% of the markets, nor would you want to trade a market that failed in most of the trend strategies. For example, cattle lost in every system, and the RNQ, REF, and MAB systems were consistently unprofitable; therefore, none of those would be good candidates for trading. The most consistent systems were CHL, PAR, MII, and DMC, each posting 8 of 12 winning markets; however, each has a noticeably different technique. T-bills present an interesting choice because the highest profit, 239.8%, occurred in one of the least reliable systems, REF; DRP, MII, or CHL are likely to be better choices because of their overall consistency.

Although this study is now dated, the trending nature of these markets remains the same. The German mark (now the euro) and T-bills are the best performers, and we would expect the same outperformance in 2011. Displaying the system results by year for all markets is another way to look at robustness. In Table 21.13, RNQ and MAB had net

TABLE 21.13 Percentage Returns by System and Year*

System	1978	1979	1980	1981	1982	1983	1984
CHL	6.1	47.9	81.6	21.8	28.3	19.9	28.0
PAR	18.5	16.4	54.5	29.9	−31.7	−43.2	−19.4
DRM	29.3	21.7	92.2	−31.6	−22.7	7.8	−7.3
RNQ	−57.2	−69.0	−9.6	−152.1	−31.2	−135.5	−100.1
DRP	34.8	63.8	88.7	31.0	−13.4	−20.0	38.5
MII	12.9	41.6	87.8	54.6	−30.7	3.1	7.1
LSO	−47.1	−4.0	39.1	−55.1	−55.1	−37.6	−59.5
REF	−13.5	23.3	53.2	−85.2	2.0	−109.6	−69.2
DMC	17.6	26.8	85.4	5.9	22.9	1.9	−1.6
DRI	−60.7	−15.4	46.3	−108.1	−114.4	−105.3	−34.0
MAB	−38.0	−44.3	−7.7	−114.4	−91.1	−81.8	−45.7
ALX	28.3	38.6	82.6	−34.4	−8.2	−2.0	−14.4
Average	**−5.8**	**12.3**	**57.8**	**−36.5**	**−28.8**	**−41.9**	**−23.1**

*Equal percentage amounts of margin are assumed to be invested in each market.

losses over all markets in all years, while CHL, MII, and DMC stand out as very consistent. The combined presentation of results across all markets and systems is one way of avoiding overfitting by looking at the big picture.

PRICE SHOCKS

Ignoring or underestimating the significance of a price shocks will result in a catastrophic loss—if not now, then sometime in the future. Price shocks are large changes in price caused by unpredictable and significant events. The impact of price shocks on historic tests can change the results from profits to losses; it varies the risk from small to extremely large. Unfortunately, not enough thought is given to how these moves affect test results and future performance. A discussion of price shocks could easily fill an entire book, but the concepts that are most important can be explained briefly. When it comes to actual trading, the difference between the expected results and actual performance (not attributed to slippage) is greatly dependent on the number of price shocks.

By its very nature, a price shock must be unexpected. However, not all similar events cause shocks. An assassination, abduction, or political coup is likely to be a surprise, while a crop freeze can be anticipated as weather turns unusually cold. Some economic reports, such as a 0.5% increase in the Producer Price Index or a jump in the balance of trade, will come as a shock, but a low carryover supply of soybeans or a tightening of the money supply after steady economic growth can be anticipated. During times when good economic news is important to support a continued bull market in equities, any disappointing surprise will cause a shock. Every word used by the Fed when stating its current policy on interest rates, following an FOMC meeting, is dissected by the traders, often resulting in large, sudden price changes. The market always tries to anticipate the results of a scheduled economic report. When it is right, prices remain calm; when it is wrong, prices react in proportion to how poorly it was anticipated.

Misinterpreting the Past

In this chapter, our interest in price shocks is how it affects test results and, consequently, the selection of parameters used for trading. At the time of a price shock, such as 9/11, we know that the event could not have been anticipated; at best we might have been lucky and held a short position in equities or a long in bonds. We would never attribute that profit to trading skill. The problem comes when testing a strategy with historic data, which always contains price shocks. The computer cannot know the difference between a profit from a consistent move in the right direction or a windfall profit from a price shock; results only show the set of parameters that had the best performance. For example, if there were 10 major price shocks in 20 years of data,

and your historic tests profited from 8 of those shocks, then you have overestimated your profits by profiting from at least 3 shocks that in all probability should have been losses. Even worse, you have *underestimated your risk* by posting profitable results that should have been losses. It is the erroneously low risk that is fatal to trading because it allows overleveraging.

Consider that a price shock only occurs because traders and investors are liquidating their positions or entering new positions in the direction of the price shock. If the shock favored lower interest rates or a stronger U.S. dollar and most traders already held those positions, then there would be little trading activity due to the news and correspondingly little price movement. The news was correctly anticipated, hence no price shock. The only time a price shock is important is when its interpretation is contrary to the direction held by traders. Unless you are a contrarian, a price shock is more likely to cause a loss than a profit.

To better understand the effects of price shocks on performance, a test was run that removed the profits and losses from 1-day moves where the current trading range was more than 5 times the average true range of the previous 10 days. Figure 21.21 shows that price shocks consistently had a negative effect on trend performance. Then choosing parameters that netted a profit on all price shocks would be unrealistically optimistic.

How can this problem be avoided? First, look at the pattern of profits and losses from a series of tests. If a 20-day and 30-day moving average each showed a maximum drawdown of $10,000, but the 25-day average only lost $5,000, assume that the 25-day results could have had the same drawdown as the nearby tests. By some quirk of timing one test was able to avoid a loss while a nearby test was not. It would be risky to assume such good fortune in trading. It is safer to average the returns of a series of tests, but use the largest drawdown.
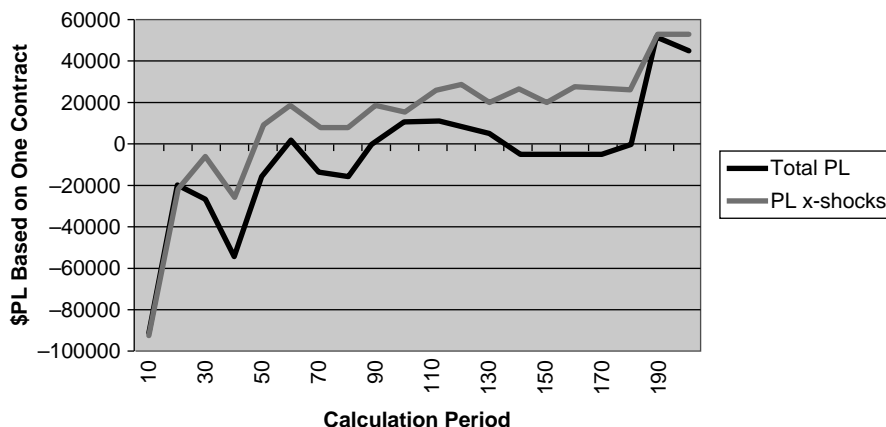
FIGURE 21.21 Price shocks consistently reduce the returns of a moving average system applied to the S&P, 1991–2009.

Create a list of major price shocks. This can be done in a spreadsheet by finding 1-day price ranges that are at least 3.5 times the size of recent average ranges. Then look at them on a chart for confirmation. Verify how many of these events produced profits or losses. If the results profited from more than 50% of the shocks, then both profits and risk are wrong. Reverse the average of the 1-day profits with the equivalent loss. Be sure only one day is changed because the system should adjust after the day of the price shock. Also check to see that there is an even distribution with regard to the size of the price shocks. Not only should there be an even number of profits and losses from price shocks, but they should be of comparable size. The system may show good results by taking losses on the minor price shocks and large profits on major ones.

Price shocks must be treated as a serious risk; otherwise, any trading account can be overleveraged and undercapitalized. This is the most common cause of catastrophic loss. A price shock that causes a windfall profit could just as easily have produced a devastating loss. The example that follows, "Anatomy of an Optimization," shows how expectations and reality are often quite different. Further examples can be found in Chapter 22.

ANATOMY OF AN OPTIMIZATION

The Iraqi invasion of Kuwait on August 6, 1990, and the U.S. retaliation on January 17, 1991, create a price pattern that causes a series of bad optimization tests; this pattern can teach an important lesson.[18] The strategy is a simple moving average applied to crude oil futures. The test period began January 2, 1990, but 100 days were needed to initialize the calculations; therefore, the first trade was entered May 24, 1990 (see Tables 21.14 through 21.16). The first test could have used more data; however, the results would have been similar.

In both Test 1 and Test 2 the moving average calculation periods selected were those that produced the highest profits. By accepting these choices the trader would have held a short position before the invasion of Kuwait, and a long position before the U.S. retaliation. In reality, there would have been large losses—the best past parameter values could not predict the price action on the next day. When a price shock is an important part of the data being tested, the best optimized result is the one that took the most profits out of that move, even at the cost of losing on most other trades. The system is not robust; it is finely tuned to one price move.

In Test 2, although there were large profits from the first price shock, the trader would have still held the wrong position when the second shock occurred. This method of testing hides the real trading risk, focuses on profits that cannot be predicted, and does it all at the expense of everyday trading profits.

---

[18]Perry Kaufman, "Price Shocks: Reevaluating Risk/Return Expectations," *Futures Industry* (June/July 1995).

TABLE 21.14 Test 1. Optimizing Crude Oil, January 2, 1990, through August 3, 1990*

Period	Net Prft	ROA	MaxDD	Trds
5	2615	82.9	−3155	11
10	6765	959.6	−705	5
15	6975	989.4	−705	3
20	4975	705.7	−705	3
25	6975	989.4	−705	3
30	4635	657.5	−705	3
35	3635	313.4	−1160	3
40	3215	276.0	−1165	3
45	3255	187.6	−1735	3
50	255	10.6	−2415	3
55	−715	−21.1	−3385	5
60	2625	156.7	−1675	5
65	2785	239.1	−1165	5
70	−385	−12.6	−3055	7
75	1955	112.7	−1735	3
80	−5040	−100.0	−5040	2
85	−5040	−100.0	−5040	2
90	−5040	−100.0	−5040	2
95	−5040	−100.0	−5040	2
100	−5040	−100.0	−5040	2

Trade Detail from Testing

Trade Date	B/S	Price	Profit/ Loss	Cum P/L
May 24, 1990	Buy	23.58		
May 25, 1990	Sell	23.55	−55	−55
Jul 11, 1990	Buy	22.49	1035	980
Aug 3, 1990	Sell	28.51	5995	6975

*Testing stopped the day before the invasion of Kuwait.
Test selects 15-day moving average.
Holding short on August 6, 1990, when Iraq invades Kuwait.

Test 3 looks back at the best results when the data includes both price shocks, August 6, 1990, and January 17, 1991. Instead of calculation periods of 15 and 20 days, selected in the first two tests, this optimization shows best returns for a 35-day calculation period. A look at the sequence of results in the optimization places the peak returns in an uncertain position between two losses, indicating instability. The trading detail shows that the system would have shorted oil on the day before the invasion of Iraq, a likely case of overfitting. The most reasonable selection would be in the area of the 60-day period, although profits are far lower, and there is still no assurance that this choice would be profitable the next time.

TABLE 21.15 Test 2. Optimizing Crude Oil, January 2, 1990, through January 16, 1991*

Period	Net Prft	ROA	MaxDD	Trds
5	–6010	–32.3	–18610	32
10	3660	22.3	–16430	20
15	12350	103.8	–11895	14
20	19810	345.4	–5735	10
25	8410	105.5	–7970	10
30	8420	86.8	–9705	8
35	18075	1348.9	–1340	5
40	11150	308.9	–3610	6
45	9335	304.1	–3070	5
50	8265	246.7	–3350	3
55	2485	55.0	–4515	7
60	9075	184.8	–4910	5
65	7155	100.4	–7130	7
70	3815	53.3	–7160	7
75	7195	83.9	–8580	9
80	–8540	–82.6	–10340	12
85	–3860	–54.2	–7125	8
90	1960	35.7	–5485	4
95	440	9.5	–4615	4
100	–3950	–85.6	–4615	2

Trade Detail from Testing

Trade Date	B/S	Price	Profit/ Loss	Cum P/L
May 24, 1990	Buy	23.58		
May 25, 1990	Sell	23.55	–55	–55
Jul 12, 1990	Buy	23.49	35	–20
Oct 19, 1990	Sell	40.10	16585	16565
Oct 30, 1990	Buy	41.23	–1155	15410
Oct 31, 1990	Sell	41.92	665	16075
Nov 19, 1990	Buy	38.39	3505	19580
Nov 21, 1990	Sell	37.30	–1115	18465
Nov 26, 1990	Buy	40.62	–3345	15120
Nov 27, 1990	Sell	40.53	–115	15005
Jan 7, 1991	Buy	35.70	4805	19810

*Testing stopped the day before the U.S. retaliation.
Test selects 20-day moving average.
Total profits are $3,225 without price shock.
Holding long when U.S. attacks Iraq.

TABLE 21.16 Test 3. Optimizing Crude Oil, January 2, 1990, through March 28, 1991

Period	Net Prft	ROA	MaxDD	Trds
5	−17255	64.0	−26975	45
10	−13195	−45.0	−29295	31
15	4480	26.6	−16840	20
20	8430	67.6	−12475	16
25	−4690	−22.6	−20725	20
30	−1395	−6.4	−21775	17
35	28335	2114.6	−1340	9
40	−360	−2.5	−14420	12
45	18245	594.3	−3070	7
50	8265	246.7	−3350	3
55	2485	55.0	−4515	7
60	9075	184.8	−4910	5
65	7155	100.4	−7130	7
70	3815	53.3	−7160	7
75	7195	83.9	−8580	9
80	−18225	−81.3	−22410	13
85	−12945	−70.4	−18390	9
90	−4065	−40.3	−10080	5
95	−8645	−56.0	−15430	5
100	−3485	−75.5	−4615	5

Trade Detail from Testing

Trade Date	B/S	Price	Profit/Loss	Cum P/L
May 24, 1990	Buy	23.58		
May 25, 1990	Sell	23.55	−55	−55
Jul 16, 1990	Buy	24.16	−635	−690
Nov 9, 1990	Sell	40.58	16395	15705
Jan 14, 1991	Buy	38.76	1795	17500
Jan 16, 1991	Sell	39.36	575	18075
Mar 8, 1991	Buy	29.89	9445	27520
Mar 13, 1991	Sell	30.89	975	28495
Mar 19, 1991	Buy	30.75	115	28610
Mar 28, 1991	Sell	30.50	−275	28335

*Test period includes both price shocks.
Test selects 35-day moving average.
Total profits are $2,495 without price shocks.
Optimization profits include both Iraqi invasion of Kuwait and the U.S. attack on Iraq.

SUMMARIZING ROBUSTNESS

We return to the concept of robustness. Earlier we listed the criteria that are important in "Searching for Robustness," the first being a sound premise. In showing test results we concluded that a large number of profitable tests was a good measurement of robustness. For the testing process, it was necessary to set performance expectations and make realistic assumptions about the parameter values that were likely to work; realistic costs also must be used. The process of testing needed to be done scientifically, using in-sample and out-of-sample data.

If we pass the hurdle of in-sample success, followed by good out-of-sample results, we have *proved the concept*. We can retest with all the data then trade, using a small commitment to begin.

Practically speaking, a *robust* trading strategy is one that produces consistently good results across a broad set of parameter values applied to many different markets tested for many years. *Robustness* is a term used to describe a system or method that works under many market conditions, one in which we have confidence. In the best scenario, a robust system is not sensitive to moderate changes in parameter values.

A strategy can also be robust if the underlying premise has a single set of rules with no parameter values and is successful across a specific set of markets during a specific time period, provided there is a large sample of trades. Ideally, if a long-term moving average system is robust, then someone unfamiliar with the system should be able to choose a calculation period between 100 and 300 days, specify a market and a time interval, and expect the system to show net profits from trading. This would be most likely if a large percentage of all the parameter values were profitable.

The highest profit or the best risk ratio is not important when looking for robustness, and focusing on these values can be distracting. Instead, it is the average values of all tests that are important, that is, the average net profit, average number of trades, average drawdown, and average profit per trade. In addition, the standard deviation of profits for all tests will give a very good idea of the consistency of returns over the range of tests.

The following is a checklist and brief explanation that outlines the steps for testing, evaluating the test results, selecting parameter values, and monitoring actual trading, consistent with the purpose of finding a robust trading strategy:[19]

1. *Deciding what to test.* Testing should be used for validation of an existing idea, not for discovery.
   a. *Is the underlying premise sound?* It should be based on careful observation of the market, or an awareness of relationships between different factors that drive the market, or between different markets themselves. It may be implemented using trends, oscillators, and other techniques, but it is trying to capture a pattern in human behavior, or an economic or fundamental process.

---

[19]A more detailed explanation of testing "robustness" can be found in Perry Kaufman, *Smarter Trading* (New York: McGraw-Hill, 1995).

  b. *Can you program all the rules?* Programming a system guarantees that all of the contingencies have been thought out. Rules can even account for exceptional situations, such as crisis intervention following a dramatic price shock, where a different set of rules are applied.

  c. *Does the strategy make sense only under certain conditions?* A strategy can apply to a smaller set of specific situations, provided there are enough cases and those patterns can be extracted from a varied set of test data. When there are fewer cases it is more important that the justification be explained in advance, based on a clear market phenomenon, and not discovered by testing.

  d. *Take a guess as to the expected results.* Based on the rules of the strategy, a rough estimate of the number of trades, percentage of profitable trades, and even the net profits should be written down before testing begins. Testing then becomes a way of confirming these estimates.

  e. *Be critical of good results.* Getting good test results in a pattern that is different from expectations may mean that the rules are specified incorrectly, or that you misunderstood the risk of the method. It is the best time to rethink the problem before moving forward. While we always try to understand and fix bad results, there is an inclination to accept good results without questioning whether they could be wrong. Getting good results for the wrong reason is still an error.

2. *Deciding how to test.* There are a number of steps necessary before the test process begins:

  a. *Choose the testing tools and methods.* Select a test platform, such as TradeStation, MetaStock, or Ninja Trading, that can speed up the test process and keep it objective. Programming in these platforms prevents the strategy from looking ahead at the data before making a decision about today, an error that can easily happen when programming in BASIC, C, or even using a spreadsheet. Many perfect strategies have been created by accidentally looking at tomorrow's prices.

  b. *Use as much of the right data as possible.* More data is best, but if there is less data available it should have a wide variety of price patterns, such as bull and bear markets, price shocks, and sustained sideways periods.

  c. *Separate the data into in-sample and out-of-sample sets*. The only way to assure that no overfitting has occurred is by reserving a substantial part of the data for validation. Once the out-of-sample data has been used, you cannot change the program.

  d. *Choose the range of parameter values that applies to the strategy.* The nature of the trading strategy, for example, a long-term or short-term breakout, will determine the parameter values that should be tested. Test only the range that makes sense for the strategy. This serves to validate the concept and avoids discovering an isolated profitable situation that can be distracting and misleading.

  e. *Decide the order in which to test the parameters.* Not all parameters should be tested at the same time. A few of the most important parameters should be tested

first to see if the basic structure is profitable. These parameters, such as the calculation period for the trend, will have a wide range of results. These results may show that faster trading cannot be profitable after transaction costs. Other parameters, such as a stop-loss, are used to refine the risk profile, and should not turn a losing method into a profitable one. They can be tested later.

f. *Be sure the parameter test values are distributed properly.* Testing equal increments of trend periods can be misleading. The difference between a 3- and 4-day moving average is 33%, while the difference between a 99- and 100-day average is only 1%. When testing periods from 1 to 100 days, in steps of 1, the average results are heavily weighted toward the long end, which has similar performance. Commercial testing platforms may not allow unequal test increments; therefore, it is necessary to weight the results in a way that compensates for the distribution.

g. *Define the success criteria.* Highest net profits are not the best measurement of success for individual tests. Risk-adjusted returns, high reliability, frequency of trading, and time in the market can be combined to give very good objective criteria.

h. *Presentation and visualization of the results.* A table of test results, in the order they are processed by the testing program, is a very difficult format for evaluation. The availability of charts via a spreadsheet allows a visual presentation of how the values of each parameter interact with all other parameter values. It becomes clear which parameters are stable and which are not performing properly. Two- and 3-dimensional contour maps and scatter diagrams can give a better view of the interaction between parameters; they also show the sensitivity of each parameter value at the same time—in easy-to-read colors.

3. *Evaluating the results.* Careful review of the test results can avoid many hours of unnecessary work later on. Looking at a good sample of individual test results (one result for each combination of parameter values), and even the individual trades of a few selected parameter values, will give insight into the good and bad aspects of a trading strategy. The statistics are not enough information to be sure that the program is performing correctly. An exceptionally large trading loss, and too many trades that ended with a stop-loss and not enough with profit-taking, will indicate a problem that cannot be seen without studying the detail.

a. *Verify a reasonable sample of calculations.* Just because calculations were done by a computer doesn't mean they are correct. Someone had to tell the computer what to do, and people make mistakes. This is the time to remember the Carpenter's Rule: *measure twice, cut once.*

b. *Were there enough trades to have reliable results?* Results that show excellent performance with only a few trades are a problem. When there is not enough data, there is no way to be sure that the results are robust.

c. *Does the trading system produce profits for most combinations of parameters?* The best assurance of robustness is when most parameters produce profitable results, even though the returns and risk vary.

    d. *Did rule changes improve overall test performance?* When a special rule is added to improve performance or to correct a specific problem, the average statistics of all tests must improve, while the standard deviation remains unchanged. Or, the standard deviation of all tests must decline while the average returns remain unchanged. Both profiles indicate a change that is a global improvement; otherwise, the changes have targeted a single problem, which is a sign of overfitting.

4. *Choosing the specific parameter values to trade.* Choosing the final values to be used in trading is a unique and difficult problem. There is no sure way of knowing if the final selection of values will perform in the future. The testing process tries to separate robust performance from success by luck or chance.

    a. *The final parameter selection can include all data.* Good test procedures call for using in-sample and out-of-sample data, but once that has been completed successfully, a final test can be run on all data. Staying as current as possible improves the chances for success.

    b. *Choose values from a broad area of success not biased by the one test that had peak profits.* Peak performance of any sort is an exception not likely to be repeated. Select the final parameter values that represent average returns in an area of consistency that is not sensitive to moderate changes in parameter values.

    c. *Profits must be distributed evenly throughout the test period.* The best parameter value choices are those that show consistency over time. The stock index futures have a tendency to be highly profitable for trending methods during the bull market of the 1990s, but have a decaying profit profile during the more recent years. A set of values that show a moderate increase in performance in recent years is more desirable that one that shows a decay in returns; however, uniformity across the entire period is best.

    d. *Verify whether a disproportionate amount of profits came from price shocks.* It is not possible to avoid large price shocks in real trading because you never know when they will come. Historic results that do not show large (negative) equity jumps due to price shocks are not realistic, and result from overfitting. The final values chosen may have produced conservative results by avoiding all of the price shocks. Even worse, those results that profited from more than 50% of the price shocks, or from the largest price shocks, have erroneously low risk. In real trading, far fewer than 50% of the price shocks will become a windfall profit.

    e. *Scale the risk to a personally acceptable risk level.* Every return is associated with a comparable level of risk. For example, the final system performance shows a risk-adjusted annualized return of 50% with an associated annualized risk of 20%, measured as 1 standard deviation of the monthly returns. Then there is a 2.5% chance that the system can lose 40% in one year. A trader that will only accept a 2.5% chance of a 20% loss will need to reduce the leverage by one-half, reducing the expected returns to 25% per annum.

f. *When there is a choice between two sets of parameter values, choose the one that causes less time in the market.* Parameter values that generate comparable returns and have less time in the market are best. Being out of the market is the only defense against price shocks. If a strategy holds a position only 60% of the time, then there is a 40% chance that a price shock can be avoided. Strategies that are out of the market most often do not reverse positions. Separating exits and entries improves slippage and liquidity.

g. *Out-of-sample performance must validate the results.* Before beginning trading, an out-of-sample test should confirm that the system can perform on data never seen before. If there is no recent out-of-sample data available, a test of a similar market or older data can be used. While this is not as good as more recent data in the tested markets, a further decay in performance when applied to a similar index would still be a strong sign that the model was overfitted.

5. *Trading and performance monitoring.* The development of a system is never complete. There is a continuous evolution in the markets, seen in the changing participation, volatility, interrelationships between markets, and the introduction of new markets. The only way to see how these changes affect a trading program is to carefully monitor the system signals and the corresponding execution prices. From this evaluation, funds can be reallocated, the method of placing orders can be changed, and better rules can be introduced. Without monitoring the trading, and comparing the results to the test expectations, there is no basis for saying that a system is performing correctly.

a. *Trade using exactly the same rules that were tested.* It is often convenient to program a simple version of an idea, but trade a more complex set of rules. Or, a new rule may be added at the end without testing, because the test process takes too long. What appears to be an inconsequential change can cause substantially different results. If the tested method and traded method are not identical, there is no basis for assessing success.

b. *Trade the same data that was used in testing.* The use of a continuous back-adjusted series for testing futures will not give you the same results as testing each contract, beginning at the time it becomes the most active delivery month. Some artificial continuous price series are smoother than real prices and appear to perform better, but are not real data; at the least, it does not include the cost of rolling contracts. Forex data is a unique problem because test data may represent the quotes of one set of banks, while the data on any quote machine will be from an entirely different set of banks. These can vary dramatically. Testing using cash FX data is not the same as futures. The best trading results come when the test data and trading data are from the same source.

c. *Monitoring the difference between the system and the actual entries and exits.* Slippage, the difference between the system price and the actual market execution, can change performance from theoretically profitable to a real trading loss. Monitoring this difference will tell you the realistic execution costs, the maximum volume that can be traded, and the best time of day to trade.

A Final Comment on Testing

Testing is the only way you can know if a strategy can be successful, but it provides no guarantee of profits. Testing can be done on many levels. It can be as simple as validating a single set of rules with no choice of parameter values, applied to past data. It can be as complex as a multidimensional analysis of results based on parameters that are so interdependent that they must all be tested simultaneously.

Testing that follows a sound procedure and a test evaluation that looks for robust solutions will improve the chances for success. Bad testing can be tragic—it is a waste of time and resources, and leads to a system that has little chance of success.

The extent of testing is a matter of judgment. Simple tests can outline the performance profile and provide an understanding of the risks and rewards. Complex tests may introduce subtle errors in the process, just as complex strategies may be flawed due to their intricacy. Begin simply, increase the complexity slowly, and stop as soon as possible.

Remember that, as hard as we try, real-time results will not be as good as historic results. Historic price patterns never repeat themselves in the same way. The best system rules and the best parameter values are the simplest and the most general. They may show lower returns and greater risk than the ideal optimized results, but they are much more likely to represent the future returns.

Other areas of interest to readers actively testing and evaluating test performance are *genetic algorithms*, a new search technique, discussed primarily in Chapters 20 and 24, and most of Chapter 23, especially the section "Measuring Return and Risk."

Practical Considerations

This chapter discusses areas of technical and mathematical analysis that are not systems by themselves yet are essential to the successful development of a system and to trading. The first part concerns the use—and the misuse—of the computer. Technical analysis and computers have become inseparable and, with the simple steps provided by development software such as *TradeStation*, *MetaStock*, and spreadsheets, it would be rare to trade a new technical system without first running it through a computer to simulate its past performance.

The second section, "Extreme Events," is the most underestimated deterrent for finding the best system, for generating realistic expectations, and the most common reason for catastrophic loss. Price shocks have been discussed in Chapter 21 ("Anatomy of an Optimization"), yet they cannot be underestimated. Unpredictable events, once they become part of price history, are often treated as though they could have been anticipated. This can inflate performance returns and minimize risk during strategy testing with serious consequences to actual trading. This section tries to explain how to close the gap between testing and reality.

"Gambling Techniques—The Theory of Runs," the third section, is actually an application of gambling techniques, primarily Martingales, to trading. What would seem more reasonable than assuming that the odds are against you in the stock or futures markets and treating it as a gambling scenario?

The final sections start by combining systems that focus on different market characteristics, and then explore the trade-offs inherent in the basic types of strategies. There is also a short analysis of trading limits in futures markets to show some of the problems and flaws of moving from a test to an operating environment. The last section reviews the similarity of systems and the impact of technical systems on the markets.

USE AND ABUSE OF THE COMPUTER

Make sure your present report system is reasonably clean and effective before you automate. Otherwise your new computer will just speed up the mess.

—Robert Townsend[1]

Computers are not a substitute for thinking. They excel in performing the same tedious task, over and over again, quickly and accurately—provided, of course, that correct information was entered. But even though technology has not yet reached the stage depicted in *Star Wars*, *Star Trek*, or even in H.G. Welles' *The War of the Worlds*, the computer has become an essential part of trading, from viewing charts to order entry to system development. This section will consider both good and bad ways to approach a computer problem.

Speed and capacity are no longer a concern when choosing a computer. Even the basic models have sufficient power to solve all but specialized problems. Storing tick data on all stocks from their inception would still present a problem, but there are servers that can easily handle that. Even laptops have storage measured in terabytes and perform multitasking. Still, buying the high-end hardware ensures some added longevity to your purchase.

During the 1980s the average life of a computer *generation* was about 10 years; however, advances in technology have reduced that to about 3 years, if you don't mind 1 year when you feel obsolete. Each year, computers become faster and more sophisticated. It is said that the capacity and speed of the computer doubles every year and a half.

Because not everyone knows how to program a computer, software companies have developed *turn-key* packages, programs that are *user-friendly*, operated by a click of the mouse and driven by menu selection. The combination of these two factors, cheap and powerful hardware and sophisticated, user-friendly software, has made individual and institutional investors dependent upon this technology. Generally, development software falls into two groups:

1. *Daily market analysis*, with price updates and orders placed at the end of the day.
2. *Intraday trading platform*, providing users with online data feeds and real-time executions.

Both types of software provide the ability to enter your trading strategies and test them on historic data. The software that operates on daily data, also called *end-of-day data*, is considered "off-line" and is used by traders who prefer to evaluate the market after it closes and prepare orders that can be executed the next day without having to constantly watch the market. The intraday trading platform, however, is much more complex and more

---

[1] Robert Townsend, *Up the Organization* (New York: Knopf, 1970), 36.

expensive. It requires active data feeds, provides real-time graphics, and generates trading signals that change as price patterns develop during the day. It can make decisions based on each price print or based on a bar size defined by the user, for example, every 30 minutes.

Acquiring Data

Daily historic data is readily available from a wide range of providers and may even come with the purchase of the software development program. A convenient list can be found on *DataVendors.com*, however, it is missing CQG and perhaps other important sources. It is common to see 25 years of historic data on all U.S. equity and futures markets, but not as common to have European and Asian data provided as well. More data and markets are always available at an added cost. Specialized data vendors, such as Dial Data and Commodity Systems Inc., provide an easy way to update data each day in a format readily used by most testing and trading platforms.

If futures data is needed, check to see if the vendor provides a continuous, back-adjusted data series, which will be necessary for testing. The types of back-adjusting were discussed in Chapter 21. You will need to understand the exact way in which data is back-adjusted so that it corresponds to the way in which you trade.

Intraday data is available from a few major vendors, such as the CQG Data Factory, Tick Data Inc., and of course, Bloomberg; however, Bloomberg does not provide an automatic download and update. Intraday trading platforms, such as TradeStation and Thomson-Reuters MetaStock, provide data feeds that are integrated with the trading strategies and can also generate orders when trading signals occur. One-stop shopping has come to trading. The Chicago Mercantile Exchange has made most of its tick data available, but it is not necessarily in a format that can be read by your trading strategy software. You may still find it easier to buy the data in the right format.

If you are using a live charting service, or a trading platform with live data, that data can be downloaded to a text file and used for historic testing. Most services will provide an "export" feature in a drop-down menu or by right-clicking on the prices that appear on the screen.

Data Accuracy

It may seem surprising, but even daily data—open, high, low, and close—can vary from one provider to another. Most often the settlement prices are the same, but the highs and lows can be off by small amounts. For most developers this is not important, but for posting the settlement value of an ETF or listed fund the differences need to be resolved each day by using multiple data sources. Occasionally, there will be a large error in the closing price that passes through, so it falls to the user to have some validity checks for unusually large moves.

Even though historic data has been available for many years, you may find short periods where the decimal point shifts within a data series. A simple visual check, displaying the historic data as a chart, will show any problems that need fixing. Smaller errors may never be found.

Combining data from different vendors may also introduce problems. One source may record a price as 135.50 and the other as 13550 or 1355.

Cleaning the Data

Tick data is notorious for including bad data with the good, simply because of the magnitude of the numbers. Errors were pervasive when futures markets were open outcry and prices were recorded by a clerk with a data entry machine, standing above the pit. Now that nearly everything is electronic, the prices are far more accurate, but not always perfect. When using high frequency data (for example, every tick), it may turn out that not all ticks are flowing through to your computer due to the speed of access. In other cases, when there is a *fast market*—extremely high volume—you will either receive the data late or miss pieces of data.

There is still good reason to check the integrity of the data. Some markets are still open outcry, and even computers make mistakes. Simple checks of daily data should include:

1. *Outliers*, those data points that are more than 4% away from the previous data (or some preset percentage).

2. *Missing dates*, more than 4 days in sequence without data should be flagged as a potential problem.

3. *Open and closing prices outside the high-low range* are not possible, and the data should be corrected or removed.

   For intraday data, some additional checks are:

4. *Frequency of data*, a comparison of the average ticks per 5-minute interval with the current frequency, looking for those times when there was no data.

5. Unusually large intraday price jumps that reverse within a few seconds indicate a data error or a large order that couldn't be absorbed easily.

One dilemma presented by cleaning data was most pronounced in the energy markets when it was only open outcry. If you developed a system based on clean data, then applied that strategy to intraday live data, you would find that there were a large number of additional signals due to bad prices. Your theoretical, profitable results would instead be filled with bad trades, netting a loss instead of the expected profit. The solution was to use the original, dirty data in the development to force the system to be less sensitive to price movement.

Combining Standard Techniques into a System

Strategy development programs provide a wide assortment of standard techniques in the form of indicators and preprogrammed systems. Choose a stochastic oscillator or an RSI, a moving average or an *N*-day breakout, and combine them in various ways with different calculation periods, such as a 10-day and 40-day smoothing, or even vary the frequency of data using 30-minute, 1-hour, and daily bars. These existing techniques can satisfy most

needs. It has already been shown that trend profits are the result of prices trending rather than the formula used to create a trendline. The trend speed is more important than the method. Adding an oscillator or momentum indicator for entry timing is also a broad concept that can be satisfied by a number of indicators. Somewhere in the combination of these techniques should be a trading strategy that matches the time frame and risk profile that you want. Before taking the path of specialized programming, there is a lot to be learned by using these standard tools.

Programming a New Idea

Computer programming talent is no longer rare, but in high demand and sometimes expensive. However, many of us have become proficient at spreadsheets and some simple programming tasks, so we have a better idea of what is good and bad. Fortunately, the data management, the complex bookkeeping, and order entry are already built in to the programmable section of trading development software. Functions such as *buy on the close* or *sell at 125 stop* allow orders to be entered as they would in actual trading. It is only necessary to code the actual strategy and the list of trades, and a performance summary will be produced automatically. Large-scale testing is also a well-used feature. Most important, the software prevents you from looking ahead at tomorrow's price in order to decide whether to buy or sell today. Throughout this book, there are many examples of strategies in both Excel and TradeStation. Anyone able to program can take the simple TradeStation instructions and change them to Basic, C, or any other language.

The prevailing philosophy about software development is to try to use as much third-party software as possible. For example, after an optimization has been run (resulting in hundreds of lines of test results, one line for each combination of parameter values), save the results to a spreadsheet. This test detail can then be sorted, organized, analyzed, and charted with all the facilities in the spreadsheet program. If a series of tests are saved, they can be compared in the spreadsheet. Results can be averaged, and the variance can be found in order to find the most robust strategy. For the more popular platforms, such as TradeStation and MetaStock, a number of users' groups, support groups, and newsletters can be of great help. Some companies provide alternate statistical measurements for your tests, while others have an add-in portfolio program. It is a good idea to check to see if these services satisfy your needs before attacking the programming yourself.

Too Much Power

It is tempting to let the computer solve the problem of finding the ideal trading system, but it can't do that. In Chapter 21, we discussed the process of finding a successful trading system beginning with a sound premise. Instead, you can throw all indicators, trend calculations, fundamental data, and rules into a big pot and let the computer find the best combination using brute force optimization, neural networks, or genetic algorithms, but by now you should know that an overfit solution does not work. More computer power does make processing faster, but is not essential to finding a successful system. There is

a good argument for saying that less power means that the solution should be reasoned out in advance to be most effective. But then, it should be reasoned out anyway.

Looking into the Past

The advantages of using as much data as possible have been discussed thoroughly—it reflects a broader variety of market conditions; therefore, results based on more data will be more robust. There is often a concern that older data does not represent current conditions. While there may be some truth in that, there is usually more to be gained with a system that can survive those years and still perform well in current markets. Unless there has been a structural change that makes the old data invalid, longer tests and more data are better. If the older years represent a period of low volatility that is not likely to be seen again, a strategy should, at the very least, not post losses for that period. Price patterns never seem to repeat themselves; however, there are characteristics of the past that continually reappear.

Paper Trading

With all good intentions, when a new strategy is developed, the plan is to hold the most recent data aside for out-of-sample testing. However, before testing is completed, that data has been used at least once, most often many times. No matter how careful you are, it is possible that the system has been overfitted to the data by fine-tuning the rules and parameters. Even if there was no conscious overfitting, the returns from real trading will be different from the tested results.

The safest alternative is a period of *paper trading*. Only by monitoring the program daily can its performance be seen under real market conditions. Even before doing that, the performance expectations must be documented. You should know the frequency of trading, holding period, percentage of profitable trades, size of profits, and size of losses in order to measure comparative results.

It is always a relief when the system posts its first profit, but it is the overall profile that is most important. When paper trading, assume that the execution is the worst price of the 5 minutes immediately following an intraday signal, or the highest price of the futures market closing range if trades are executed *on the close*. A touch of reality can point out problems before they happen and save a lot of money. Don't start trading a new system without monitoring its performance first. After that, start with a small amount.

It Should Work, but It Doesn't

It was hard work, and all of the steps were followed as conscientiously as possible. The rules were logical, and well-defined, and the strategy was tested on enough data. Your effort has resulted in proving one of four things:

1. The strategy is profitable.
2. The strategy is not profitable.

   3. The strategy cannot be implemented effectively.

   4. The whole process takes too long or is too expensive.

   Of course, a successful outcome is the best result, but it is not the most likely result from the first attempt. If you are convinced that the method should work but the results were not good enough, each step in the process will need to be reviewed. Trading strategies should be built slowly; each piece must be proved to be sound before going on to the next step. You should never have gotten to the end without knowing that the final product was going to be profitable. If the performance deteriorated as new rules and features were added, you should have stopped until the problems were resolved.

Vertical or Integrated Solutions

There are two philosophies for system development, *vertical*, using individual building blocks, one upon the other, and *fully-integrated*, similar to a puzzle box where all pieces fit together in just one special way. The fully-integrated method takes advantage of the interrelationship between features, such as the calculation speed of a trend and a filter that blocks entering a new trade and forces the liquidation of the current trade when volatility is too high. It is possible that, as the filter blocks more trades, the trend speed will change to be profitable in markets with less volatility. Therefore, the two parameters, trend speed and filter threshold, are interdependent and must be tested together.

   Using the vertical method, the primary rule is tested first and must be profitable without any other features. In this case, the trend is tested and found to be successful when the calculation period is above 40 days. Next, the volatility threshold is tested with the selected trend period. The volatility threshold is therefore dependent upon a specific trend. It becomes clear that the volatility feature works or does not work.

   What happens when the vertical solution shows that volatility does not help, but there is a very successful integrated solution? This will happen when the trend period is not profitable on its own, but is profitable when selected trades are made based on the volatility filter. To some analysts, this is a more sophisticated solution; to others, it is overfit. The answer can be found by tabulating the percentage of profitable test combinations, a measure used in Chapter 21 that was called *robustness*. If the percentage of profitable tests of the combined parameters is greater than the percentage of successful tests using only the trend, then the volatility feature is an improvement.

Transparent or Complex Solutions

A *transparent strategy* is one that is easy to understand. If it was capitalizing on a chart pattern, it could be explained by showing that pattern, or combination of patterns, on a chart. If a trading signal was generated, another trader who knew the rules could look at the prices on a screen and confirm that a signal should have occurred at that point.

   A complex system is not transparent. A chart pattern that is filtered by price volatility and volume, then confirmed by the pattern of other stocks in the same sector as well as the S&P 500 index, would be difficult to calculate manually. If the underlying premise

is a trend based on a portion of a trigonometric curve, it is not clear that anyone could intuitively know that a trade should have occurred on any specific price bar.

Complex solutions may be wrong without anyone knowing. They are more difficult to create, especially difficult to debug, and cannot be visually confirmed by studying the signals on a chart. Acceptance is based on historic testing and profitability. A complex solution can be successful, but it is a very difficult road to follow, and difficult to fix when it goes wrong.

Isolating the Problems

A few of the more common problems that cause disappointing test results are:

1. The rules were interpreted incorrectly. When the rules were programmed, choices were made in order to satisfy a rule that was not definitive. For example, the rule could have been buy only if confirmed by volume. When that rule was programmed, it was interpreted as buy only if today's volume is greater than yesterday's volume. Because volume is erratic, the statement should have been buy only if today's volume is greater than the average volume of the past 10 days.

2. The techniques did not work for the parameter value range that was tested. The original concept was to find a medium-length trend to avoid very long holding periods that have high risk, or very short-term trading that is too intense. However, none of the parameter values in the range from 15 to 70 days worked and the stop-loss of 1% to 3% did not seem to be effective.

   Systems do not always work as expected. It turns out that medium-speed trends have highly variable holding periods; that is, sometimes the price move is sustained for 1 to 2 weeks, other times 2 months. The same trend period cannot capture both. Negative results are important feedback. They help create a better strategy before it is too late. After all, if the answers were known in advance, this would all be unnecessary.

   Instead of finding a single, medium-speed trend, combine a longer trend with the medium one. The longer one will find the fundamental direction, dictated by economic policy; the medium calculation period will capture shorter moves within that trend, at the same time avoiding the losing scenario of trading against the fundamental direction. Take the time to understand why two trends worked when one did not. This is an opportunity to increase your understanding of how the market works and find a successful strategy at the same time.

3. *It worked in some markets but not in others*. Some markets have clearer trends, while others have more noise. Not all markets will work with the same strategy; however, all markets within the same sector should work. Don't expect Raytheon and IBM to trade the same way, or Intel and General Electric, but two wireless telephone companies or two auto manufacturers should produce similar results. In futures, the fixed-income markets show the longest trends, while the equity index markets have the most noise. You should be concerned when similar markets perform differently, but expect different markets to have unique characteristics.

Just Because It Doesn't Work in Practice
Doesn't Mean It Won't Work in Theory

If the more general issues are not the problem, then you need to look deeper.

Did you make assumptions? When the initial idea was verified, was it probably done with selected data, representing markets that were most likely to be successful. It may work during a particular type of market, but fail during other times. Try to isolate the same type of price patterns in other markets by filtering periods of high or low volatility or isolating intervals of high correlation.

Ignoring a few large losses. Long-Term Capital made the mistake of rationalizing a few large losses, convincing themselves that the combination of events that caused the largest equity drawdowns during historic testing would not occur again. If those losses were removed, then the expected risk was much lower, trading could be leveraged, and profits would be increased proportionally. They were right and wrong. The same events did not happen again, but different circumstances resulted in losses just as large. Because of their high leverage and massive positions, they shook the entire financial system when they fell. If anything, history does not show enough risk. The increased correlation between markets, and greater participation, will result in faster, larger, and more volatile price swings with less protection from diversification. Large historic losses should not be ignored, but accepted as a likely but infrequent event. There is more about this important problem in the next section, "Extreme Events."

Misjudging market volatility. If the year was 1996 and the equity markets were moving steadily higher, a newly developed system trading the S&P or NASDAQ 100 could be very profitable. As time advances towards 2000, the system continues to perform in the same bull market, but volatility increases significantly. If the system trades the same number of NASDAQ contracts in 2000 as it did in 1996, then it is exposed to 6 times the volatility as prices moved from 600 to 4,000. Individual trade risk must account for changing volatility by varying the stop-loss and the position size. In a diversified portfolio that included soybeans and NASDAQ futures, soybeans remained within a narrow range of volatility, while NASDAQ volatility kept growing. Did the historic tests account for this dramatic imbalance over time?

Failing to recognize an evolving marketplace. Not only will volatility increase with price, but price patterns will change with growing market participation. Increased volume from a broad set of participants introduces more noise into short-term price movement. Economic policy still creates the major trends; however, theoretical market efficiency does not stop the increase in erratic price movement caused by large and small traders acting with similar or opposite strategies. The net effect is that a trend system will take longer to recognize a trend change. It will enter later and exit later, resulting in a reduction in net profits. Over a long test interval, this will show a noticeable decay in performance.

Taking Control

It is not necessary to create or acquire a system that is exceptionally complex in order to be successful. Simple systems, such as long-term trends or a dual moving average crossover, can be profitable as well. A simple trend approach, applied to the S&P or NASDAQ, would have gotten you out of the long positions after the 2000 tech bubble peak, and well before the largest losses from the subprime crisis in 2008. While most financial managers and brokers were encouraging clients to hold their positions, that it was just a temporary setback, the trend had turned down. The problem with advice based on market "know-how" is that there is no specific point that would cause a change. It is only clear that positions should have been exited well after the fact, when it is too late. At the very least, even a passive investor needs to recognize that the money is more important to him or her than it is to the broker. Being proactive can cause a small opportunity to be missed but can also avoid a big loss.

EXTREME EVENTS

Price shocks are the most likely reason for catastrophic loss. Although very large price shocks are infrequent events, experiencing only one is enough. Price shocks cause the largest gap between expectations (formed by historic testing) and real trading results. The implications of price shocks in testing were discussed in Chapter 21.

A *price shock* is an unexpected, unpredictable event. Price shocks are a normal part of price movement and happen often. Most are the result of economic reports that differ from expectations, such as the Fed raising rates by ½% when only ¼% was anticipated, unemployment rising when it was expected to fall, or a drug company's latest miracle cure being denied by the FDA. But the largest price shocks are unexpected events with global impact, such as the terrorist attack of 9/11/2001. A significant unexpected political event, such as the abduction of Gorbachev in 1991, caused havoc in the foreign exchange markets for 3 days. The collapse of Lehman Brothers may have been a shock, but the subprime crises evolved over days and weeks. The key word to remember is *unexpected*.

Money Moves the Markets

As technology has made communication instantaneous, everyone seems to know of a price shock at the same time. The world watched while the second airplane flew into the World Trade Center. Markets plummeted until trading was halted in the United States. Drops in the S&P 500 were mirrored in the equity markets of every country.

During a price shock, diversification is ineffective. Markets all move together, or all reverse together. While the equity markets dropped, fixed income prices skyrocketed in a *flight to safety*. Both of those price moves are easily explained. However, other markets, such as soybeans and cocoa, uniformly reversed direction because of large-scale

liquidation. The fundamentals are no longer important—investors are simply cashing out of all positions. It is the money that moves the market during a price shock. It was exactly the same scenario in the subprime crisis of 2008. Correlations went to 1, not because those markets had anything fundamentally similar, but because investors were moving money out as quickly as possible. Markets that are entirely unrelated to the events will become volatile in what appears to be a sympathetic move. This usually compounds the trading risk while it presents great opportunities for profits after the price shock has been fully absorbed.

Fooling Yourself with Hindsight

If an event was expected, then it would not cause a price change, let alone a price shock. No one can profit from a price shock by clever planning, but only by luck. We can never assume that an open trade would be on the right side of a price shock in more than 50% of those events, and it may be far fewer. Unfortunately, backtesting a trading system using historic data does not attempt to identify price shocks, but treats them as normal, everyday events. The final selection is often the parameter values that produced the best returns or return ratio, without concern about specific profits that may have been the result of shocks. Regardless of the way the best performance was judged—higher profits, lower risk, least sensitive, or a combination of statistical values—you can be sure that selection has been the greatest beneficiary of these unpredictable price shocks. It goes a long way toward explaining why out-of-sample and real trading is never as profitable as the simulated test results.

While there are no other alternatives for validating a proposed trading strategy, backtesting is weakest because of its handling of price shocks. In an attempt to understand the enemy, four important prices shocks since 1990 were:

1. *The abduction of Gorbachev.* A particularly difficult price shock on August 16, 1991, was followed by an opposite shock two days later. Russian Premier Gorbachev was abducted and then surprisingly released (see Figure 22.1). European currencies dropped sharply against the dollar because of anticipated disruption of Russian trade. Most financial markets were affected at the same time, as were the energy markets. Many currency traders profited from this shock because they were holding short positions over the weekend when Gorbachev was abducted on a Sunday. They continued to hold shorts looking for a further decline, but gave back their profits when prices abruptly reversed on Wednesday.

2. *Labour's loss in the British election, January 1992.* Politics holds many surprises, but the victory of John Major's Conservative party in 1992 was unexpected (see Figure 22.2). After a 2-month uptrend, prices gapped lower and posted a daily range of nearly £10, close to 6%. The low was made on the fifth day after the election results and was followed by a 3-week recovery that retraced 90% of the fall. Unlike the Gorbachev abduction, this price shock had limited effect on other markets.

FIGURE 22.1 The Gorbachev abduction caused a double price shock, with most impact on the European currencies, such as the US$/German mark.

Source: Chart created with TradeStation® by Omega Research, Inc.
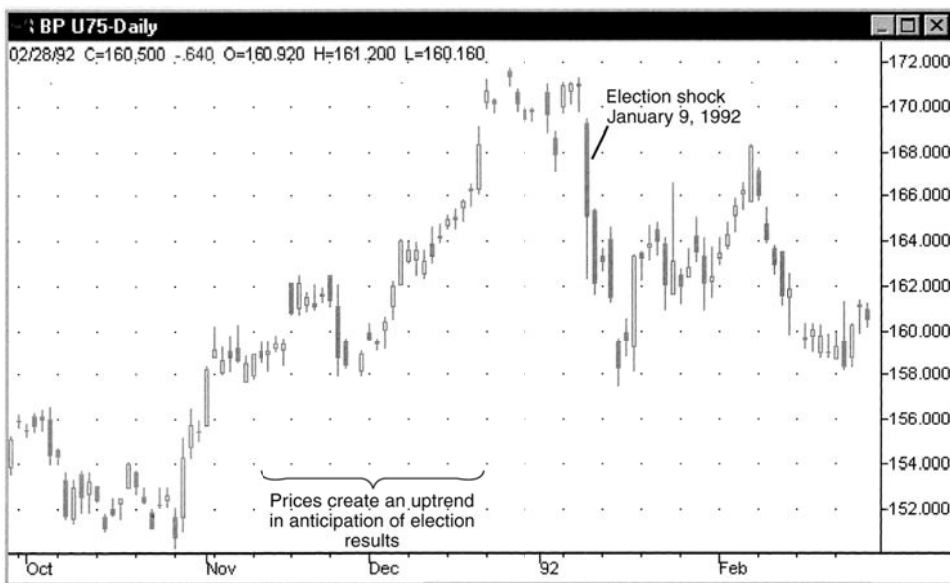
FIGURE 22.2 Labour's victory in the British election of 1992 represents a more common shock.

Source: Chart created with TradeStation® by Omega Research, Inc.
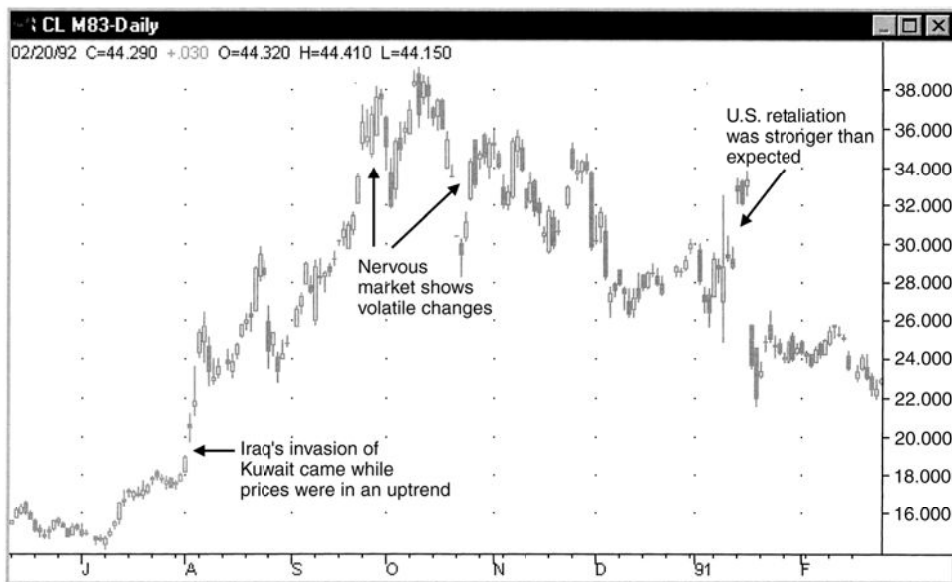
FIGURE 22.3 Crude oil prices from Iraq's invasion of Kuwait through the beginning of the Gulf War, August 1990 to February 1991.

Source: Chart created with TradeStation® by Omega Research, Inc.

3. *The Gulf War, August 1991 through January 1992.* Oil prices are notorious for volatility, but the Gulf War marked a period of unprecedented price shocks driven by news (Figure 22.3). Because Iraq had been moving troops near the Kuwait border, Iraqi intentions were not a surprise. In Figure 22.3 prices were rising from the second week of July as news of Iraq's plans circulated in the marketplace. The threat of an oil supply disruption and a failing diplomatic solution resulted in a rise of $4, or 28%. Therefore, many systematic traders and commercials would have been long on August 7, 1990, when Iraq invaded Kuwait. During the next months, oil prices show a number of gaps and wide-ranging days, fueled by news and uncertainty. The most dramatic price shock was on the day of the U.S. invasion of Kuwait, January 17, 1991. Futures prices opened near their high of $30.80/bbl and by the close of trading had fallen to $20.50. The first cash trade in oil, after the invasion was known, exchanged hands at about $40/bbl. Preceding the U.S. invasion, oil prices were exceptionally volatile but showed an upwards trend. There would have been no trend system holding a short position ahead of the $10 drop in futures prices.

4. *The terrorist attack on September 11, 2001.* This generation will remember where they were on 9/11, just as the last generation remembers where they were when John F. Kennedy was assassinated. In an unprecedented move by the exchanges, all trades executed in the morning before the attack were canceled (Figure 22.4). After the S&P close on September 10 at 1101.4, prices of futures opened 6 days later
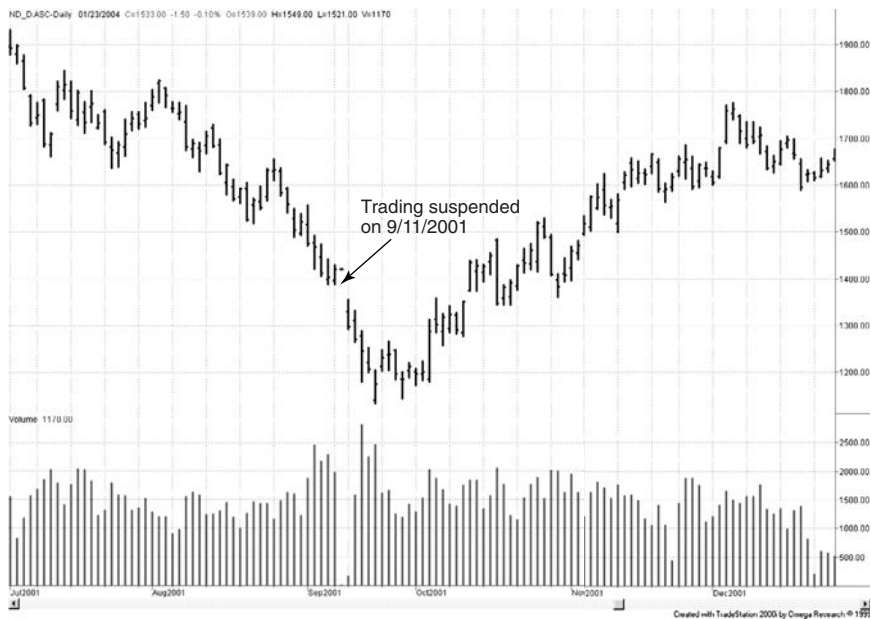
FIGURE 22.4   The price shock of 9/11/2001.

on September 17 at 1041.4, 5.4% lower, posting the low for the move of 939.0 on
September 21, after 5 days of steady declines.

The price shock of 9/11 favored the trend followers. The two months prior to 9/11
show an increasing bearish move, with the last two weeks steadily declining. The shock
itself, the difference between the close of 9/11 and the open of 9/17 was not as severe as
the other examples. The surprising 42-point rally on September 28 may have produced
more losses for systematic traders than the actual price shock. The steady rise that
followed during the next few months is similar to the recovery after the October 1987
plunge. Traders see the extreme lows as the farthest prices can be pushed, and most
sentiment turns to the other direction.

The interesting elements in the four examples are:

- *Price shocks do not always generate losses for trend traders.* The British election and
  the U.S. invasion of Kuwait caused prices to change direction. The Iraq invasion of
  Kuwait and 9/11 favored current trend positions. The Gorbachev abduction was good
  on one day and bad two days later, or just the opposite, depending on your position.
- *Prices change direction after the shock has been absorbed into the market.* A price
  shock pushes prices to an extreme, and after the market has absorbed the news, it
  is most likely to move in the other direction for a longer period. If a trading account
  was capitalized adequately to absorb these shocks, a crisis management strategy

may be able to recover much of the loss or even net profits from the second phase of the reaction.

- *Volatility always increases*. If the initial shock resulted in a windfall profit, reducing the size of the position, or even exiting the entire position, is a reasonable way to deal with exceptional risk.

Not all news is a complete surprise. In the agricultural markets, a crop freeze in orange juice or coffee is often anticipated by a change in weather. Before a freeze can occur, the temperature must drop, and below-freezing temperatures must be sustained in order to cause damage. This weather change causes processors to protect themselves by buying forward contracts or futures. In turn, prices move up in anticipation of problems. However, on January 9, 2012, the USDA announced that it was looking into the possible contamination of Brazilian orange juice from a chemical fungicide. Brazil provides $^1/_3$ of all the orange juice in the U.S. market. Prices jumped 16% in 2 days, then reversed 15% on the following 2 days when the report turned out to be questionable.

Identifying Price Shocks

If we cannot predict price shocks, and they seriously impact our historic tests, then we must understand their effects. In Chapter 21, it was shown that price shocks always added profits to historic performance, but in actual trading, this is not likely to be true. This was done by identifying a price shock as a 1-day trading range five times the size of the previous trading ranges, as measured using the average true range. By accumulating the net change relative to your trading position, on those days identified as a price shock, you can see the impact on performance.

Crisis Management

One aspect of crisis management is having sufficient financial reserves to survive an extreme, correlated price reversal, if not a worst-case scenario. For most trading programs, holding enough reserves aside to weather a worst-case scenario will lower the annualized rate of return to a level below the return on government bonds. Nevertheless, the potential for a large loss is real and must be taken into account. That means some tactical method must be used.

In the previous examples of well-known price shocks, prices reversed direction within a few days and continued in that direction for some time. If the original position was held after its initial loss, then part of that loss would have been recovered. The strategy that evolves is

1. If there is a windfall profit from the price shock, exit the position.
2. If there is a large loss from the position, hold the position expecting a reversal. This may take a few days and may increase the risk.
3. If there are adequate reserves, the position can be increased on peak volume or when volatility declines at a later point.

Holding a position during a price shock seems very risky; however, the trade has already seen most of the risk by the time the first order to exit could be placed. In Chapter 14, we showed the price patterns that included the price shock and the following reaction. In most cases, the reactions were orderly and the increased risk from holding the trade for a few days is small in proportion. During Gorbachev's abduction, a short position should have been liquidated, getting the low price, while a long position would have begun its recovery immediately, reaching new high profits 13 days later. The British election had two legs down and only recovered to unchanged after 3 weeks. During the U.S. invasion of Kuwait, a long position, liquidated immediately, would have been highly profitable; a short position held for only a few hours more would have been even more profitable. Shorts held on 9/11 should have been liquidated immediately, even though they would not have captured the full profits, but longs would have suffered additional losses for the next 5 days. In 2012, the orange juice price shock reversed the entire move within 4 days. All of these situations require careful monitoring, but all have the same characteristic that the shock of the news carries prices to extreme levels, followed by a substantial reversal.

System Disconnect during a Crisis

A large price shock can create a disconnect between a trading system and the market. This is particularly true for trending methods. During 1987, the stock market began an accelerated decline in mid-October, leading to a 2-day drop of large proportion on very high volatility and confirmed by a peak in volume. Figure 22.5 shows how 50-, 100-, and
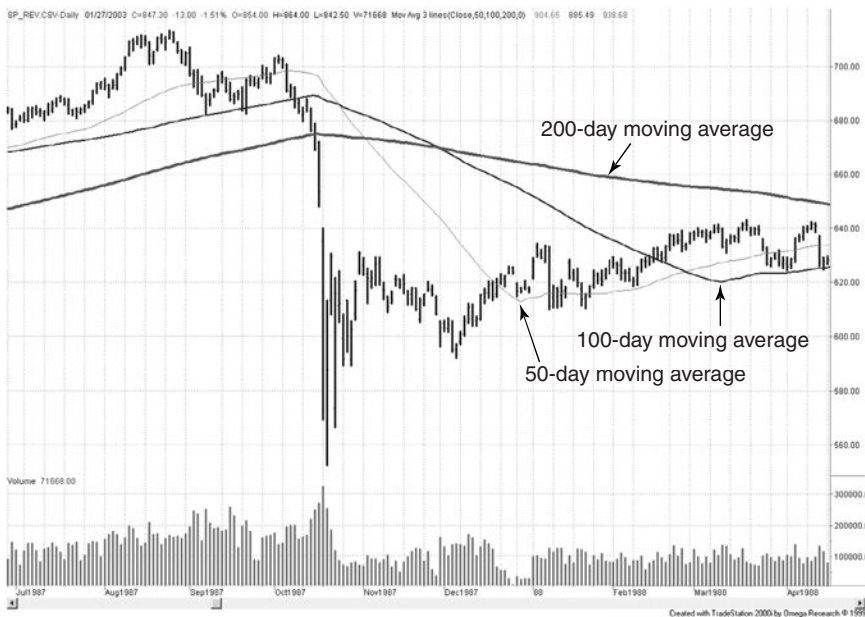
FIGURE 22.5   S&P 500 futures. Moving average trendlines become out-of-phase during a crisis.

200-day moving averages reacted to that move. All three moving averages turned down on the day of the largest drop, and following a trend signal to sell would have captured the lowest price of the move. Switching to crisis management rules, where shorts are liquidated and longs are held, would have been a far better strategy.

Equally important is the direction of the price movement compared to the direction of the three trendlines after the low price. The fastest, 50-day moving average turned down, crossing the prices that were moving higher in mid-December, and finally turning up in late December. The 100-day moving average turned up in the following March, and the benchmark 200-day moving average turned up in August of 1998. It seems clear that these moving averages are out of phase with the price movement after the shock. The moving averages are not reflecting the market trend; they are catching up to an event long past. Dogmatically holding a short position from mid-October onward because the trendline is down puts the trader out-of-phase with reality.

The collapse of the world stock markets in the fall of 2008 was not technically a price shock, and there was no one day that could be attributed to the decline. The news came out slowly, steadily, and clearly in favor of an economic disaster. Figure 22.6 shows that any long-term trend follower would have been short ahead of the decline, which, on this portion of the price movement, was a drop of 55% in the S&P futures. Here again there is a disconnect in the moving averages, but the sustained concerns over the future of the economy allowed the moving averages to catch up to the lower prices and trigger an upturn at a reasonable point. Had prices turned up sooner, the trends would still be declining. Because of this fortunate pattern, and their ability to hold shorts, trend followers trading futures posted their best year in the history of that industry.
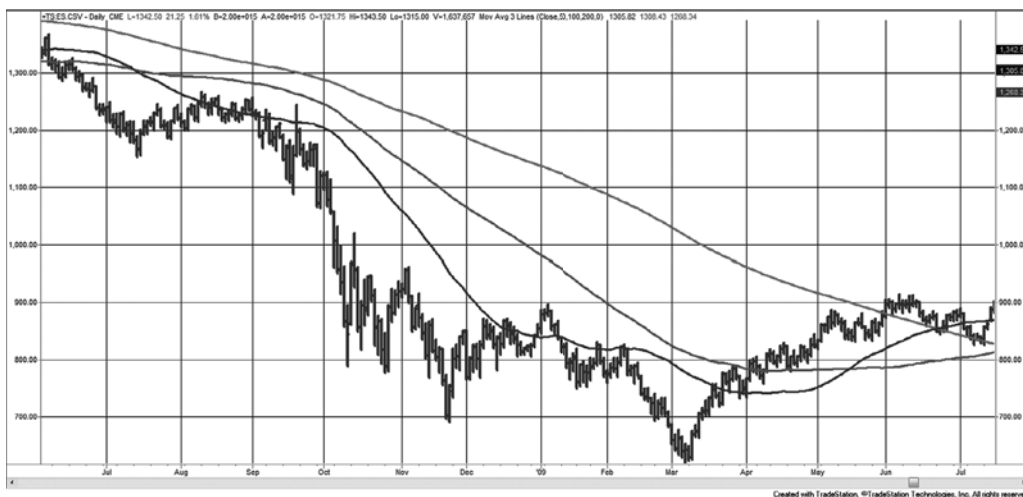
FIGURE 22.6 The S&P decline due to the subprime crisis generated large profits for trend followers who were already short. They were fortunate that the depressed prices allowed the trends to catch up to lower prices.

Crisis management must be invoked to correct unusual situations that are not handled by the main strategy. These events can be identified using a technical measure of a price shock. The new rules can be well-defined in advance or they can be handled by a team of competent managers. They must recognize that, as in the Crash of 1987, the trendlines are no longer useful. Once the extreme period has passed, which can be determined by a drop in volatility, the trends can be reinitialized at the new price levels and follow the new price direction.

GAMBLING TECHNIQUES—THE THEORY OF RUNS

The application of gambling theory to highly leveraged or short-term trading satisfies two important conditions. First, it presumes no statistical advantage in the occurrence of profits and losses but concerns itself instead with the probable patterns. Viewed simply as an up or down price move, or a trading profit or loss, each daily change could be treated as an occurrence of red or black in roulette. Second, gambling theory stresses money management under adverse conditions. A successful professional gambler and an active trader must be highly disciplined and conserve capital. This section will look at a gambler's approach to money management and risk, using the *Theory of Runs.*

If the assumption is made that the profit or loss of each trade is unrelated to the previous trade, or that each successive price has an equal chance of going up or down, the situation closely resembles roulette. In Monte Carlo, the roulette wheel has 37 compartments: 18 black, 18 red, and 1 white, assuring a loss of 2.7% in the same way that transaction costs are a handicap to trading. Continuously betting on only red or black will break even, less 2.7%, over long betting periods. The only way to change the odds is by money management—varying the size of the bets. Although the variables in trading are more complex, we will look at this first.

The most well-known method for winning in a gambling situation is based on the probability of sequences occurring, the Theory of Runs. On each spin of the wheel, the likelihood of the same color (red or black) reoccurring is 50%, or 1:2 (1 out of 2, ignoring the white slot). In order to define a run of 3 reds, it is necessary to show a black on each side, that is,

$$\textit{Black-Red-Red-Red-Black}$$

otherwise the run may not be complete. If each color has a 50% chance of occurring, the probability of a run of 3 reds is

$$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \left(\frac{1}{2}\right)^5 = 0.03125$$

or about 3%. One run of 3 can be expected for every 32 spins of the wheel (called *coups).* Extending that to runs of n consecutive reds gives $(1/2)^{n+2}$. For 256 coups, which is both

a power of two and the approximate number of trading days in a year, there are the following possibilities for runs of red:

Run of Length	Probability of Occurrence	Expected Number of Occurrences	Total Appearances of Red
1	$\frac{1}{8}$	32	32
2	$\frac{1}{18}$	16	32
3	$\frac{1}{32}$	8	24
4	$\frac{1}{64}$	4	16
5	$\frac{1}{128}$	2	10
6	$\frac{1}{256}$	1	6
			Total appearances: 120

There is less than a 1% chance that a run greater than 6 will occur, and that only 4 will occur out of 10,000 spins. Notice, however, that the total appearances of red is only 120, short by 8. These 8 appearances could increase any of the runs from 1 through 6, or become a run of 7 or 8. The likelihood of a run greater than 6 is calculated using a geometric progression to get the sum of all probabilities greater than 6, or $(1/2)^{n+2}$, where $256 \geq n \geq 7$:

$$P = (1/2)^9 + (1/2)^{10} + \cdots + (1/2)^{256}$$

$$= \frac{(1/2)^9}{1-(1/2)} = \frac{1}{256}$$

There is a single chance that there will be a run greater than 6 in 256 tries. The average length of a run greater than 6 turns out to be 8, based on the decreasing probability of occurrences of longer runs. The average length of all runs greater than n will be $n + 2$. That makes the table of runs complete, with the number of occurrences of red equal to 128.

Each time you double the number of spins, you increase the length of the longest run by 1, so that for 512 spins there should be a run of 7 and for 1024 spins, a run of 8.

Martingales and Anti-Martingales

The classic gambling application of the Theory of Runs is called *Martingales*. In the simple version of this approach, an initial bet is doubled each time a loss occurs; whenever there is a win, the gambler nets one unit of gain and the betting begins again at the initial size. To demonstrate how this works, it is necessary to use a table of uniform random numbers (which can be found in Appendix 1). These numbers vary from 0 through 99999, but should be interpreted from 0.0 to 1.0. If you are using Excel, then select the *uniform random numbers* which return values equally across the range from 0.0 to 1.0. Let all the random numbers from 0 to less than 0.5, be assigned to red and larger values to black. Figure 22.7, read left to right, where open squares are red and solid squares are black, shows the first 257 assignments according to the Appendix 1. Assuming that we bet on black, losses will depend on the longest run of red. We must decide the size of the initial
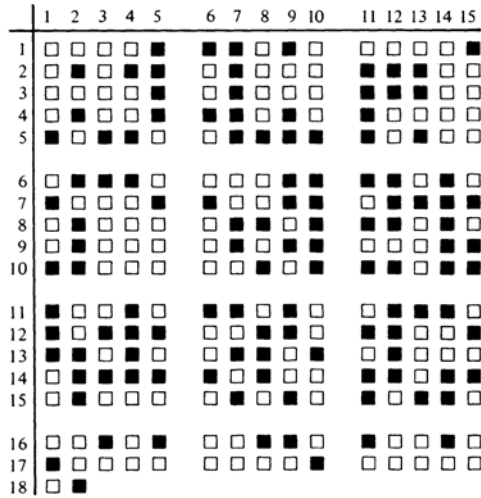
FIGURE 22.7 Sequence of random numbers representing occurrences of red and black.

bet in advance, which should be as large as possible and still withstand the longest run of red that is likely to occur. By using the results from the analysis of the length of runs, we find that for every 256 coups it is likely that only 1 run greater than 6 will occur, and that run would most probably be 8 in length. The probability of a run of 9 is $(1/2)^{11}$, or one in 1024. In 256 coups, the odds are about 3 to 1 against a run of 9 occurring.

Having decided that capitalization must withstand a run of 8, we calculate that a bet of $1 doubled 8 times is $128. Divide the maximum amount of money to be risked by 128 and the result is the size of the initial bet. If the available stake is $1,000, then when divided by 128 gives $7.8125, which must be rounded down to $7; therefore, on the eighth consecutive occurrence of red, the bet will be $897. Counting the occurrences of runs on the simulated roulette table (Table 22.1), it is interesting to see that a run of 8 (but not 7) appears.

TABLE 22.1 Simulated Runs

Red	Black
35 runs of 1	35 runs of 1
12 runs of 2	14 runs of 2
7 runs of 3	11 runs of 3
7 runs of 4	3 runs of 4
2 runs of 5	2 runs of 5
2 runs of 6	
1 run of 8	
138 total occurrences	118 total occurrences

	1	2	3	4	5		6	7	8	9	10		11	12	13	14	15
1	1	2	4	8	[16]		[1]	[1]	1	[2]	1		2	4	8	16	[32]
2	1	[2]	1	[2]	[1]		1	[2]	1	2	4		[8]	[1]	[1]	1	2
3	4	8	16	32	[64]		1	[2]	1	2	4		[8]	[1]	[1]	1	2
4	4	[8]	1	2	[4]		[1]	[1]	1	[2]	1		[2]	1	2	4	8
5	[16]	1	[2]	[1]	1		2	[4]	[1]	[1]	[1]		[1]	1	[2]	1	2

FIGURE 22.8   Betting pattern.

The results are within expectations, and no runs greater than 8 occurred in either red or black. For the purposes of this example, the betting would then proceed as shown in Figure 22.8, using an initial bet of $7. Every occurrence of black is a winner. Although as much as $897 may be won on a single black coup after a run of 8, each sequence of runs nets only the initial bet of $1 because of the accumulated losses during that sequence. For the 256 coups shown in Table 22.2, there were a total of 65 distinct runs resulting in a profit of $455, a return of 45.5% on capital. Instead of using Martingales, if we had used the basic strategy of betting equal amounts on black to win on every coup, we would have lost $140 on a $7 bet. The Martingales method therefore has a good chance of winning a moderate amount.[2] In a casino, the drawback is the occasional run of 9. The house sets the maximum bet for the table, which always prevents doubling down after a run of 8.

Anti-Martingales

The *anti-Martingales* approach offers a smaller chance of winning a large amount; it is exactly the opposite of applying Martingales. Instead of doubling each losing bet, the winners are doubled until a goal is reached. Because there is an excellent chance of 1 run of 6 in 256 coups and a similar chance of a longer run, this method wins if the long run occurs in the first half of the number of coups to be played (256 in this case). Once a run of 6, occurs you must immediately stop playing.

We already know that a run of 6 returns $32 on a bet of $1, and a run of 8 nets $128. This method would have lost if it had been applied to black in the test sequence; because there were 138 red coups with no black runs greater than 5, there would have been a loss of $138. If the bet had been on red, looking for a run of 6, there would have been three wins, each for $94, and a loss of $118 for that many appearances of black. Waiting for a run of 8 would have won $128 and lost $117 on black by stopping right after the win. The success of anti-Martingales depends on how soon the long run appears. In 4,096 coups, a run of 11 will occur once, returning $1,024 on a bet of $1. In the same 4,096 spins, there will be 2,048 losses, showing that if the long run happens in the middle of play, the method breaks even; if it occurs sooner, you win.

---

[2] The Martingales approach is likely to work if the player could withstand adverse runs of 11, but casinos tend to limit bets to amounts significantly less than 210 times the minimum bet.

The Theory of Runs Applied to Trading

Before applying either Martingales or anti-Martingales to the markets, it must be determined whether the movement of prices up or down is as uniform as in roulette. A simple test was performed on 15 diverse futures markets from 2000 through mid-2011 and the combined results of all up and down runs are shown in Table 22.2. The top section of the table shows the total days in each market, followed by the total number of up and down days and the average upwards and average downward price change. This allows us to understand the bias in the data. The lower part shows the differences between the normal occurrence of sequences and those that actually occurred in each market.

   The two most extreme differences are the DAX and Eurodollar interest rates. The DAX shows an unusually high number of 1-day reversals, with most of the other combinations less. The Eurodollars are the opposite, showing far fewer reversals for sequences 1 through 5 (with an exception at 3), adding those differences to the longer sequences to create a fat tail. The differences between these two markets and the normal sequences, as well as the average of all markets, are shown in Figure 22.9. For both the DAX and the Eurodollar, and for all markets on average, there are fewer runs of 4 and 5 days.
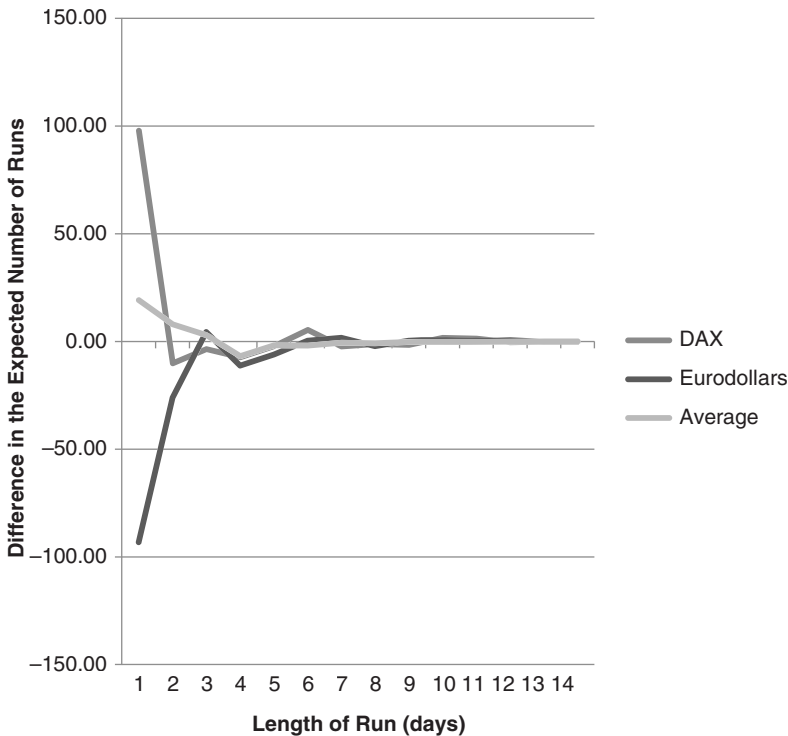
FIGURE 22.9   The two most extreme markets that differ from the normal sequence of runs, as well as the average of 15 diverse markets.

TABLE 22.2 Expected Occurrences of Ups and Downs for a Diverse Set of Markets

	Normal	S&P	NASDAQ	DAX	30-Year Bond	Euro-bund	Euro-dollars	EURUSD	USDJPY	Crude	Heating Oil	Natural Gas	Gold	Copper	Wheat	Cotton
Total days	2632.6	2617	2634	2678	2577	2657	2443	2622	2616	2632	2634	2568	2612	2614	2604	2623
Days up		1425	1397	1422	1371	1392	1301	1373	1324	1363	1320	1226	1417	1349	1242	1309
Days down		1192	1237	1256	1206	1265	1142	1249	1292	1269	1314	1342	1195	1265	1362	1314
Avg up		9.175	18.404	53.833	0.579	0.307	0.0467	0.0061	0.0047	0.9899	0.0286	0.1552	5.881	3.312	8.6457	0.9803
Avg down		−11.062	−21.136	−61.362	−0.613	−0.306	−0.0434	−0.0063	−0.0047	−1.0419	−0.0279	−0.1563	−6.017	−3.252	−8.1863	−0.9799

Length of Run	Normal	S&P	NASDAQ	DAX	30-Yr Bonds	Euro-bund	Euro-dollars	EURUSD	USDJPY	Crude	Heating Oil	Natural Gas	Gold	Copper	Wheat	Cotton	Average
1	658.15	22.85	−16.15	97.85	−39.15	16.85	−93.15	27.85	−5.15	28.85	45.85	26.85	36.85	69.85	56.85	9.85	19.12
2	329.08	−6.07	−16.08	−10.08	28.93	32.93	−26.08	12.93	32.93	−3.07	13.93	−9.07	23.93	46.93	−10.08	6.93	7.93
3	164.54	1.46	1.46	−3.54	1.46	−6.54	4.46	5.46	−0.54	16.46	19.46	−7.54	−0.54	−23.54	15.46	20.46	2.93
4	82.27	5.73	2.73	−7.27	−3.27	−13.27	−11.27	−19.27	2.73	−12.27	−22.27	−1.27	−13.27	−2.27	2.73	−11.27	−6.87
5	41.13	−6.13	5.87	−2.13	−5.13	7.87	−6.13	−3.13	−3.13	3.87	−2.13	−2.13	−5.13	−2.13	−10.13	1.87	−1.87
6	20.57	−0.57	−5.57	5.43	−3.57	−2.57	0.43	4.43	−5.57	−1.57	−4.57	4.43	−4.57	−3.57	−7.57	−4.57	−1.97
7	10.28	1.72	4.72	−2.28	1.72	−3.28	1.72	−2.28	−5.28	−3.28	−2.28	−1.28	−2.28	−2.28	4.72	0.72	−0.62
8	5.14	0.86	−0.14	−1.14	−1.14	0.86	−2.14	−2.14	−2.14	−2.14	0.86	−3.14	4.86	−3.14	−1.14	−2.14	−0.88
9	2.57	−2.57	−2.57	−1.57	−0.57	2.43	0.43	3.43	0.43	2.43	1.43	−1.57	−0.57	−1.57	−2.57	−0.57	−0.24
10	1.29	−1.29	1.71	1.71	−0.29	−0.29	0.71	−1.29	−0.29	−0.29	0.71	−1.29	−0.29	−1.29	−1.29	−1.29	−0.29
11	0.64	0.36	−0.64	1.36	−0.64	−0.64	0.36	−0.64	0.36	−0.64	−0.64	−0.64	−0.64	0.36	−0.64	−0.64	−0.24
12	0.32	−0.32	−0.32	−0.32	−0.32	−0.32	0.68	−0.32	−0.32	−0.32	−0.32	0.68	−0.32	−0.32	−0.32	0.68	−0.12
13	0.16	−0.16	0.84	−0.16	−0.16	−0.16	−0.16	0.84	−0.16	−0.16	−0.16	−0.16	−0.16	−0.16	−0.16	−0.16	−0.03
14	0.08	−0.08	0.92	−0.08	−0.08	−0.08	−0.08	−0.08	−0.08	−0.08	−0.08	−0.08	−0.08	−0.08	−0.08	−0.08	−0.01

Earlier tests of these runs had shown that there were fewer short-term sequences and most markets had shifted those differences to the fat tail. That no longer seems to be the case. The DAX, with many more 1-day reversals, exhibits an increasing amount of noise compared to expectations, while the trend characteristics of the Eurodollars have fewer short-term changes but do not increase the tail by an obvious amount.

One way of incorporating these results into a trading program is to treat the higher or lower frequency of short-term reversals as a way to improve entry prices:

- After a new trend signal, wait for the DAX to reverse before entering the new trade.
- When there is a new trend signal in the Eurodollars, enter immediately.

The most interesting applications are in the betting strategies of the Martingales and anti-Martingales methods, applied to futures markets. The following sections show two possibilities.

Martingales within a Trend

By combining the Theory of Runs with the direction of the trend, the chances of being on the correct side of the longest run are increased, and the size of the price move in the direction of the trend may also tend to be larger. Long-term trends will offer the best chance of identifying the direction of prices. We can create rules to take advantage of this:

1. If the trend has just turned up, based on a moving average, enter a long position with 1 contract.
2. If the trend is still up and the price closes lower, double the size of the long position.
3. If the trend is up and the price closes higher, remove all positions in excess of the original 1 contract.
4. If the trend turns down, exit all longs and sell short 1 contract.
5. If the trend is down and prices close higher, double the size of the short position.
6. If the trend is down and prices close lower, cover all short positions in excess of the original 1 contract.

Because no one has unlimited money and it is possible for price to go the wrong way for up to 8 days, possibly 9 over a long test period, we need to cap the number of times we can double down. For this test, we will start with 1 contract and allow up to 8 contracts. At that point we will hold the current position until prices move in our favor, then reduce the size back to 1, the initial amount. To find the results, the program *TSM Martingales with Trend* was used, available on the Companion Website.

Two markets were tested, Eurodollar interest rates and the S&P, because they represent the extremes of a very trending and very noisy market. Only a 60-day trend was used. Because there will be a lot of entering and exiting positions, a $25 round-turn cost was charged to avoid unrealistic results. Table 22.3 gives a comparison of three results for Eurodollars. The first row are the results of a straight 60-day moving average, using

| **TABLE 22.3** | Comparison of Moving Average Results with Martingales with and without Cap, Applied to Eurodollar Rates, 2000 through Mid-2011 |

	Net PL	ASD PL	Ratio
MA 60	16375	2456	6.67
MA 60 x 8	131000	19648	6.67
Martingales	549518	67262	8.17
Martingales cap 8	529243	63898	8.28

1 contract only, where the trading signal is based on the direction of the trendline. In the second row the same method used 8 contracts for each trade. The third row gives the results of the Martingales strategy with rules 1–5 above, and the last row is the Martingales strategy with a cap of 8 contracts.

The second row, using 8 contracts, is needed to compare the last row, where the contract size varies from 1 to 8. In that case, there must always be enough money in the account to trade 8 contracts, even when 1 is the current position size. In Table 22.3, the first column of numbers, *Net PL*, is the total accumulated profits for the strategy, the next column is the annualized standard deviation of the daily profit/loss changes, and the last column is the ratio of the two previous columns, essentially an information ratio. For our purposes, the information ratio is the best measure of results and shows that the capped version exceeds all others. The *Net PL* of rows 2, 3, and 4 are shown in Figure 22.10. The Martingales results might be good in this one case, where the Eurodollars are extremely trending, but what about the highly noisy S&P futures?
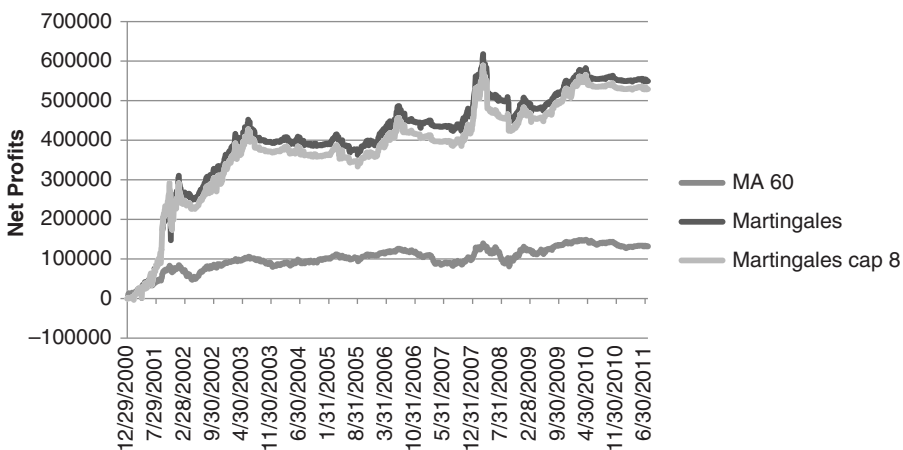
FIGURE 22.10 Comparison of Net PL for a 60-day moving average system trading 8 contracts, a Martingales approach, and a Martingale capped at 8 contracts, Eurodollars 2000 to mid-2011.

TABLE 22.4	Martingales Comparison Applied to the *emini* S&P, 2000 through Mid-2011		
	Net PL	**ASD PL**	**Ratio**
MA 60	13363	11371	1.18
MA 60 x 8	106900	90966	1.18
Martingales no cap	1365112	355856	3.84
Martingales cap 8	252687	243369	1.04

One would think that a noisy market would benefit from doubling the position size on a negative move because there would be a high likelihood of a reversal and a fast recovery of the losses. But the S&P noise hurts the ability to identify the trend with any consistency, so that even a 60-day trend filter does not improve the chances of being right by much. Table 22.4 shows that the S&P performance ratios are much lower than those of Eurodollars and the capped Martingales, the only one we could really trade, underperforms all the other combinations. But then, none of the combinations are particularly good. A chart of the net profits is given in Figure 22.11.

Anti-Martingales

Applying anti-Martingales to a trading situation presents more of a problem. The approach calls for adding to the position size when we are profitable and resetting the size to some nominal amount when we have a loss. Unless you trade a very large account, the size of the profits from a normal gain may not be enough to add one more contract. Then, if you have a small losing day, you would need to reduce the position size fractionally, but there is no benchmark size if you have been increasing the position size as the equity grows.

The only practical applications seems to be doubling the number of contracts each time prices move in a profitable direction, then reverting to 1 contract after 3 or more profitable
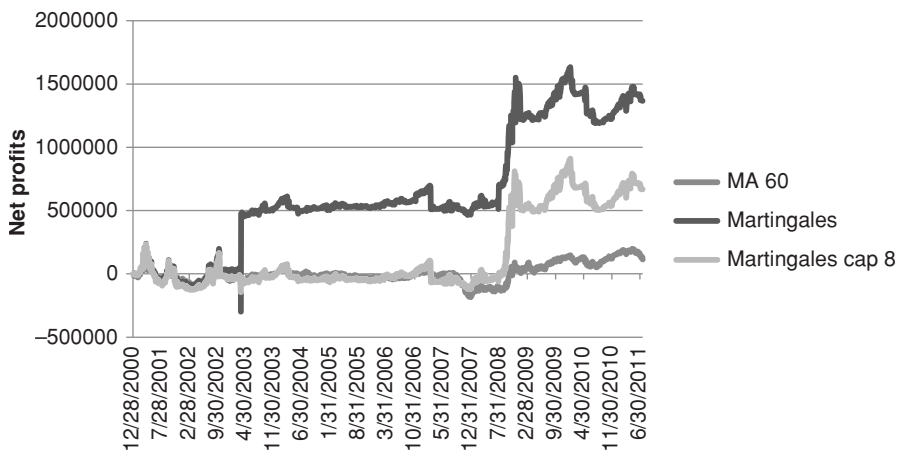
FIGURE 22.11 Comparison of Martingales net profit for the *emini* S&P futures, 2000 through mid-2011.

TABLE 22.5 Anti-Martingales for Eurodollar Futures, 2000 through mid-2011, Filtered with a 60-Day Moving Average

	Net PL	ASD PL	Ratio
MA 60	16375	2456	6.67
MA 60 x 8	131000	19648	6.67
Anti-Martingales no cap	−16443	535752	−0.03
Anti-Martingales cap 2	916037	83639	10.95
Anti-Martingales cap 4	815743	78707	10.36
Anti-Martingales cap 8	742093	75667	9.81

days, much like the gambling approach. Success depends on having a large number of runs in the profitable direction. The original anti-Martingales method calls for leveraging up in anticipation of 6 moves in the profitable direction then moving to another table, hoping that the run of 6 came in the first 50% of the 100 flips. There is no counterpart for stopping if the market goes the right direction, so we need to play for the continuous sequences of runs. Table 22.5 shows the results of Eurodollar tests using a 60-day moving average to decide the direction of leveraging. The rules were identical to the Martingales approach except that, when a profitable day occurred, the position size was doubled, and when a losing day occurred the position was reduced to the initial size. The cap limits the size of the position, so that a cap of 4 allowed the position to be doubled twice.

In Table 22.5, row 2 is simply the *Net PL* and volatility of row 1 multiplied by 8 to allow better comparison with the cap 8 row. The ratio does not change by scaling all positions. The *no cap* result is most interesting because it posts a net loss. That can be seen clearly in Figure 22.12.
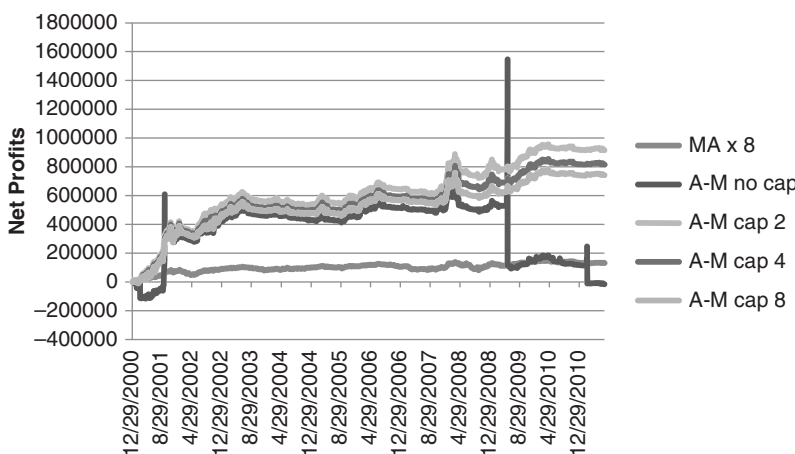
FIGURE 22.12 Comparison of Anti-Martingales method with different caps, applied to Eurodollar futures, 2000 through mid-2011.

Except for the *no cap* case, allowing the position to double for up to three times beats the simple moving average method. The *no cap* method continues to double the position until there is a loss, which then occurs while holding the largest position. There are three cases where this happens, once where the net return is still profitable and two times where the system takes a huge loss. On the other hand, the frequency of two positive days in a row is high, so that adding extra risk paid off.

A similar case was run for the *e*mini S&P, but the results were poor. Because of its high level of noise, there were more losses when positions were doubled, resulting in worse results than the simple moving average. This method can be tested using *TSM Anti-Martingales with Trend*, available on the Companion Website.

Fractional Martingales

The simple Martingales and anti-Martingales just discussed offer one possibility for applying gambling theory to position management but always double the position size. If a large account is being traded, it may be better to use the actual profits to decide the number of contracts for anti-Martingales. Another approach would be to scale down the risk by trading 1.5 times the previous position size rather than twice. This was the approach taken by Furguson,[3] and should give the same relative improvement over the basic trend method.

Delayed Countertrend Entry into a Run

For Martingales, rather than doubling the size on the first down day in an uptrend, waiting until prices posted two or three lower closes should increase the chances of a move back in the trend direction. Although there are fewer opportunities, the chances of a continued move in the same direction quickly become smaller.

There has always been conflicting interpretation about the chances of a continued price move following a series of moves in that direction. In a purely random event, there is always a 50% chance of prices moving in either direction; however, the market is not a random event, and behavioral finance advocates will argue that there is enough predictability to be profitable. This premise can be substantiated by measuring the distribution and size of the average up and down moves within a long-term trend.

Gustafson[4] studied the S&P from 1988 to 1995 and created a trading strategy that entered a long position after prices closed lower 4 days in a row, exited after an upwards run of 5 days, and always exited after 8 days. Admittedly, this method is only long and takes advantage of the upwards bias in the stock market, which was particularly strong in the mid- and late-1990s. The results for the in-sample period had 78% of the trades profitable, a $35,030 net profit and a $5,847 maximum drawdown. Applying this to data

---

[3] James William Furguson, "Martingales," *Technical Analysis of Stocks & Commodities* (February 1990), and "Reverse Martingales," *Technical Analysis of Stocks & Commodities* (March 1990).

[4] Gordon Gustafson, "Price Persistency," *Technical Analysis of Stocks & Commodities* (January 2002).

from 1982 through mid-2001, the results showed 76% reliability, profits of $348,460, and a drawdown of $56,370, showing that this pattern was successful out-of-sample, but still in the same bull market. The rules are:

1. If $C_t > C_{t-1}$ then $up = up + 1$ else $up = 0$
2. If $C_t < C_{t-1}$ then $down = down + 1$ else $down = 0$
3. If $down = 4$ then buy on next open
4. If $up = 5$ then $exit\ longs$ on next open
5. If $days\text{-}in\text{-}trade = 8$ then $exit\ longs$ on next open

SELECTIVE TRADING

It may be unrealistic to think that you can always follow a technical, fully automatic system without manual intervention. Overriding system signals may occur for reasons of sound risk management during a crisis when you decide that the system is not programmed to deal with the current volatility, that a more conservative action is the best business policy, or that the market has temporarily disconnected from the strategy.

The beginning system trader often has concerns that the impersonal buy and sell orders will not react properly to the underlying forces that are driving prices. This can cause the trader to override the system signals. For example, if the program is geared to take profits after a move of 2%, the trader might judge that the current price move is going to continue and opt for holding the trade rather than taking profits. The application of this special *intuitive* filter generally favors holding a position through ups and downs until it shows a profit (if ever), rather than applying consistently good risk management. Judgment tends to interpret all news in favor of holding the position in anticipation of a profit someday. An experienced trader has learned to control these urges to override the system, not to take a loss personally, to keep losses small, to establish and stay with a trading plan, and to filter out unsupported opinions. Trading is a business, and business principles must take precedence over emotions.

Filtering Out the Losing Trades

An investor once told a trading advisor that, "I would be happy to take smaller profits if you can eliminate the losing trades." If that were only possible, there would be many more rich investors. Nevertheless, a lot of effort goes into an attempt to automatically select the profitable trades, or reduce the number of losing trades.

Two approaches to this are most common. The use of multiple indicators, such as two moving averages, or more than one oscillator, such as a stochastic or RSI. Or, signals can be confirmed by another market, or by seasonality, or weekly patterns. Given the power of computers, some combination will produce the appearance of a very, very good system. However, we should know by now that this is just an exercise in overfitting. The

simple solution, even with risk that is larger than we want, is most often the best solution. There is no way to filter out all the losing trades and leave only the profitable ones.

There are many trading systems that can be bought that promise profits. Those that do not disclose their rules are called a *black box*. Trying to incorporate a black-box approach to improve the performance of your own program is not a good idea. Without knowing everything about the other system, there may be duplicate techniques—a fast trend, a 3-day cycle, or an increase in position size and risk. Although a correlation analysis of the black-box approach and your system may show that the methods work well together, there are still too many uncertainties. The unknown system may have been overfitted and will not show its real risk, or it may have a particular weakness to a market pattern that has not occurred during your testing period. You must understand every method that you trade, know how the parameter values were selected, and verify its results. It's your money and your risk.

SYSTEM TRADE-OFFS

Although traders are continually searching for a system that has no losses, they are realistic enough to know that most profit, loss, and other characteristics of a trading system are dependent on one another. Just as in gambling theory, a system can be structured to have frequent small losses and occasional large profits—or frequent small profits with some large losses. It is simply not possible to create a system with such small losses that these losses are essentially zero without increasing the frequency of the losses to the point where profits rarely occur. In short-term or high-leveraged trading, the *cost of doing business*, transaction costs, prevents such an approach from working. Even "risk-free" arbitrage turns out not to be risk free.

The basic principles of systematic trading require that the trader choose between the frequency of profits to losses, the relative size of those profits and losses, and the number of opportunities that will be available to trade. These interrelationships will be different for various types of systems.

Trend-Following Systems

The relative risks and rewards of trend-following systems change based on the selection of the calculation period. Using a simple moving average system applied to closing prices, the example in Figure 22.13 shows that a faster moving average will stay closer to the current price than a slower moving average. The maximum risk is measured from the current price to the corresponding moving average value.

The other changing pattern is that longer trends capture more of the fat tail. In fact, a fast trend cannot capture the fat tail at all, although the average profit should still be larger than the average loss. Basic trend following will always have many more small losses compared to fewer larger profits. At the same time, the risk, due to the lag, increases with the trend period.
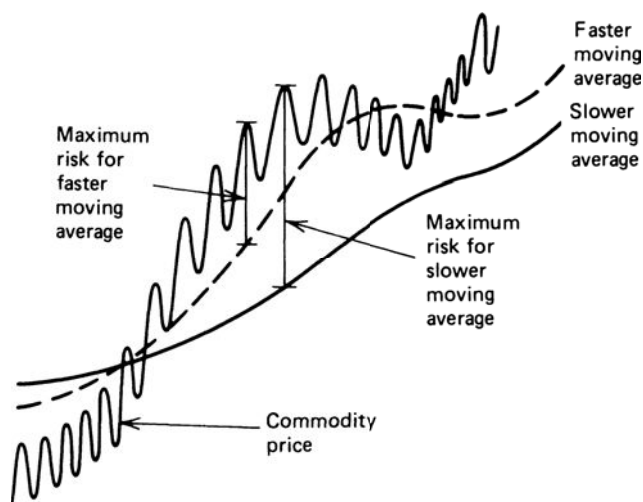
FIGURE 22.13 Relative risk of a moving average system.

The risk of a single trade being small or large is an incomplete measure of trading risk. Because there are more losses than profits in normal trend-following systems, it is important to consider the sequence of losses as part of the ultimate risk of trading. It is the aggregation of these losses that provides a reasonable comparison in determining whether a slow or fast moving average system has the best profit-to-loss ratio.

The attraction of a trend system is that it cuts losses short. This is called *conservation of capital.* On the other hand, it is unrealistic to expect a smoothing technique to produce more than 40–45% profitable trades; 30–35% is normal.

Mean-Reverting Systems

Mean-reverting systems are those that take advantage of nondirectional price movement or a fundamental divergence between two markets, selling when prices move higher and buying when they move lower. The belief is that prices will return to the mean. The risk-and-reward profile of a mean-reverting method is different from that of a trend-following approach. Although the size of the profits and losses relates to the holding period of the trade in a manner similar to trend following, positions must be set in advance of an expected price reversal, or while two markets are diverging. This introduces more risk and makes the success of the method dependent upon good timing. When mean-reverting methods are discussed here, those techniques that require a confirmation—that is, they wait until prices reverse direction—are not included. Waiting until prices turn introduces a trend element.

Due to the greater risk of a mean-reverting system, its objectives are generally well-defined. In a trend system, you let the profits run until prices reverse and the trend changes direction; however, most mean-reverting methods rely on an overbought/oversold indicator, contrary opinion, or divergence, and set goals for exiting when that condition disappears or becomes neutral. This results in a performance profile that limits profits;

therefore, those profits must be more frequent. To accomplish results with limited but more frequent profits, you cannot also limit the size of the losses.

Trade-Offs

The performance profile goes together with particular trading styles. Ideally, you would want very large profits, high reliability, and small losses. Unfortunately, these combinations are not possible. The realistic trade-offs that are associated with the two primary systematic strategies are:

Trend Following	Mean Reversion
Many small losses and a few large profits	Many small profits and a few large losses
Lower average risk per trade	Higher average risk per trade
Improve reliability with larger risk	Reduce number of trades to lower risk

The last trade-off for trend following refers to longer calculation periods being generally more successful, but at the cost of larger lags. For mean reversion, making the entry threshold further away means lower risk once the trade is entered, but far fewer opportunities.

An example of mean-reverting trade-offs are shown in Figure 22.14, where prices are moving sideways, with no apparent trend bias. This can also be viewed as a price series with the trend removed. Assuming that there is a normal distribution of prices, there is a clustering near the center and less frequent peaks and valleys as prices move further away from the middle. If high values are treated as overbought and low values as oversold, the area above and below the center line can be divided into three zones, *slightly* (S), *moderately* (M), or *extremely* (E) overbought or oversold.

In Figure 22.14 there are more opportunities to buy a *slightly oversold* market and sell a *slightly overbought* market, the points where prices cross the innermost set of horizontal lines. Profits are taken when prices return to the mean, the center horizontal line. All prices must pass through the area closest to the center before reaching a more overbought or oversold condition; therefore, the *opportunities* must decrease as profit goals increase.
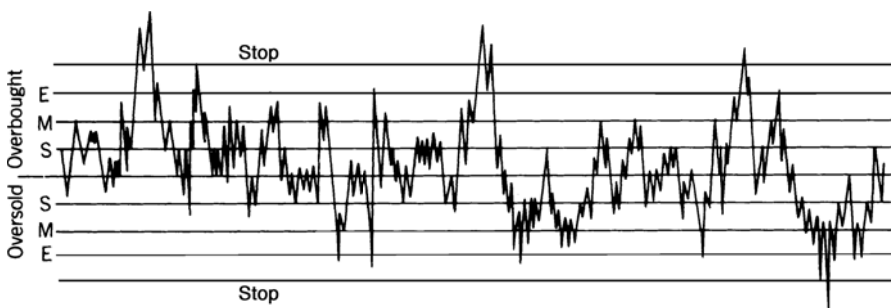
FIGURE 22.14   Entry point alternatives for a mean-reverting strategy.
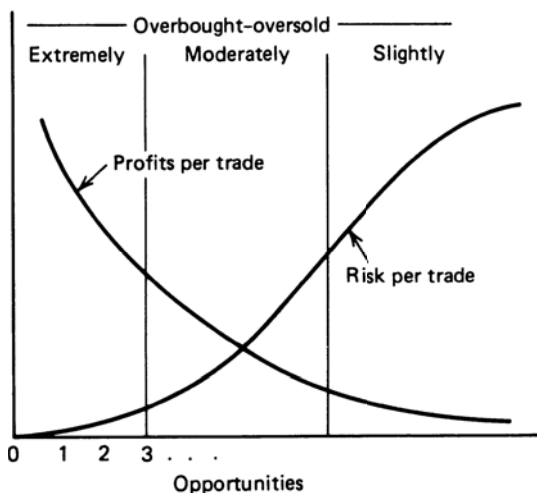
FIGURE 22.15 Relationship of profits to risk per trade based on opportunities.

But what about the risk? If trades were entered at the *S* lines, and the *M* lines were used as a stop-loss, then almost one-half the trades would be stopped out, and the size of the losses would equal the size of the profits. Because price changes are not normally distributed, there are actually fewer losses than profits using this method; however, the net returns would not be impressive. If the entry were at the *S* lines and the exit at the outermost *Stop* lines, then most trades would be profitable, but the losing trades would be large. The net of profits and losses would be better because relatively few trades reach the stop. If trades are entered at the *E* lines and exited at the *Stop* lines, then profits would be larger than losses and risk greatly reduced. However, there would be very few opportunities and a larger percentage of them would be stopped out. Theoretically, the minimum risk occurs if a trade can be entered at the absolute extreme points, an unlikely possibility, and there would be only one trade. Figure 22.15 shows the relationships that exist for a standard mean reversion system.

All systems have trade-offs. In order to reduce the risk, there must be more small losses or fewer opportunities—whether a trending or mean-reverting strategy. A trend-follower looking for the big move must also take a big risk and many small losses. In the development of a system or trading philosophy, each person must settle on the combination of risk, reward, and opportunity that best suits him or her. There is no *best* combination.

The Commodex Method of Combining Elements

The *Commodex* system was first presented in 1959 and is still actively run by second-generation Philip Gotthelf.[5] Commodex combines the components most acceptable to

---

[5] More on Commodex can be found at their website, Commodex.com.

the experienced trader in a unique weighting method. It includes moving averages, price momentum, and open interest to calculate a trend index.

The most interesting aspect of the Commodex system is its ranking process, intended to produce a relative strength value for each trend. Using 10- and 20-day moving averages, the system scores the current market performance to establish the value of the trending component. Bullish and bearish values are calculated by looking at three situations independently: the simple moving average signals derived from both the long- and short-term trends, and the double moving average signal generated by combining the two trends. The two techniques of single and double moving averages are treated exactly as covered in Chapter 8. The most important of the three factors is the long-term trend, second is the short-term trend, and last the relative position of the fast to the slow moving average. The highest-ranking upward-moving trend occurs when current prices are above both the faster and slower moving averages, and the faster average is below the slower. The opposite positions would result in the strongest downward-trending component. Trends are considered neutral if the most important element, the long-term moving average, conflicts with the other two factors.

The rate of change in open interest (for futures) is considered a reinforcement of the trend. Using a concept different from the usual charting techniques, Commodex considers it a bullish sign if there is increasing *growth momentum* in open interest combined with rising prices. The growth momentum is the difference between the rate of increase of the open interest and the 20-day moving average of the open interest. Confirming a bull move with rising prices and rising open interest is a classic concept of charting. The bear move is confirmed with rising open interest and falling prices. Commodex also considers a drop in open interest along with falling prices to be a bullish factor. The movement of volume is treated in the same manner as open interest and can confirm a bull or bear trend. An increasing volume momentum with rising prices is support of an upward move; other combinations indicate bear trends.

Added together, the signals are ranked from a *strong buy* to a *strong sell* with lesser degrees in between. The system must be given credit for the quantification and balancing of these elements, which are generally treated as highly interpretive charting techniques. The rules for applying the daily signals to trading combine both the individual strength of the signal with the movement of the Commodex trend index. The trend index itself acts as an overbought/oversold indicator, encouraging profits at specified levels and considering a position reversal in more extreme situations; stops are placed using the 20-day moving average with predetermined band penetrations. Additional objectives are based on the profits accrued from a trade, with a 50% return justifying a protective stop on part of the positions, and a 100% profit requiring the liquidation of half the position. These money-management concepts are an important aspect of any system and tend to round out Commodex.

Combining Trends and Trading Ranges

Price movement is classified as being in either a trend or a trading range, and most traders decide that one or the other of those regimes will be their sole focus. Trading ranges
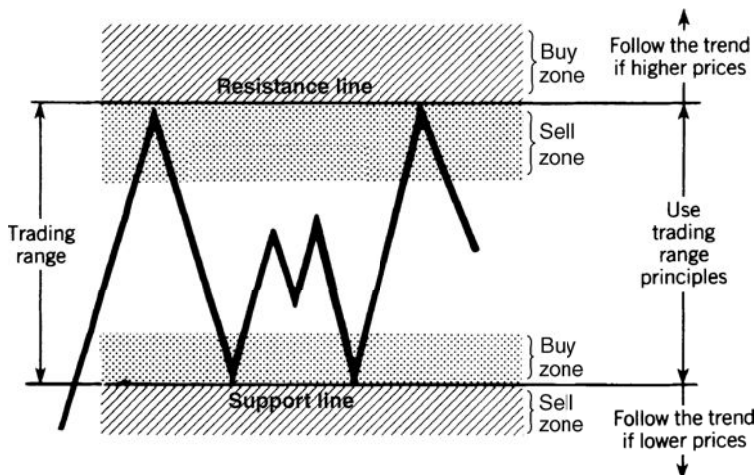
FIGURE 22.16 Combining trends and trading ranges.

and trends take turns dominating price movement, with the trading range most common, but those conclusions depend on your time horizon. In 2011, the interest rates continued to post historically low yields, while the EURUSD moved between 1.23 and 1.45. Whenever the S&P breaks above a recent high or below a previous low, analysts are quick to point to the next level of resistance or support as the price target. They imply that prices will move to those levels then hesitate or reverse direction. Because this is a well-known concept, we can put rules on combining the trend direction with clear support and resistance levels. These are shown in Figure 22.16.

The exact rules for combining the two techniques depend on the size of the trading range, the speed of the trend, and whether the support and resistance levels have been established over time. Recent range formations offer less of an obstacle than those that have survived for some time. If the range is narrow, a moving average buy or sell signal invariably occurs at the resistance or support level and is met with an initial reversal. For a larger range, a medium-speed moving average signal may occur closer to the center of the range and allow some opportunity for trend profit before resistance is encountered. For very wide ranges, there may be ample latitude for a moving average to signal entry and close out without the interference of support and resistance levels. One method of combining the two techniques is to use whichever signal is generated, regardless of the system. The following combinations of short position entry and exit would be used for medium to wide ranges:

1. Enter a new short position when the moving average turns down.
2. Enter a new short position when prices enter the sell zone below the resistance level.
3. Enter a new short position when prices fall below the support area, then close out the short position when the moving average turns up.

4. Close out a short position when prices enter the buy zone above the support level.

5. Close out a short position when prices break through a resistance level moving up.

The same rules apply in reverse for long positions. The advantage of this filtering method is that new short trades are not entered at support levels, which causes immediate losses or prolonged trades with little return.

TRADING LIMITS AND DISCONNECTED MARKETS

Trading limits, which were unique to futures markets, have infiltrated the stock market. On February 18, 2011, the SEC adopted a *limit up/limit down rule* that halts trading for 5 minutes in any single stock or ETF if that market moves more than 10% in 5 minutes. The intention is to control excess volatility; however, it would not stop an event such as the *flash crash* that occurred on May 6, 2010. There is some doubt that this rule will accomplish its objective.

Curbs on excess price movement in futures, which can be highly leveraged, were intended to allow those traders on the losing side to meet margin calls in an orderly manner. Rules require that accounts must have a minimum deposit equal to 75% of the initial margin; otherwise, positions are liquidated by the brokerage firm. Clients have up to 24 hours to meet margin calls. Limits also prevent overreaction to false news, so that a price shock based on a rumored assassination would be corrected by the next day's opening if not true, and an extreme move that could ruin many investors would be avoided.

While futures contracts are halted at the limit, the cash market is not. This can create a severe disconnect between those two markets. Soft commodities, such as coffee, sugar, and cocoa, are trading in New York as well as in London. Traditionally, New York has had trading limits and London has not, presenting interesting dynamics in extreme cases.

Rules for limits have changed over the years. At first there was a simple limit of 20¢ per bushel for soybeans; then, it was changed to expanding limits, 20¢ the first day, 30¢, then 40¢, then no limits. As of 2008, after a very volatile period, soybean limits were raised to 70¢ and corn to 30¢; there are no longer expanding limits. Trading simply stops at the limit and continues again the next day. All agricultural futures markets have trading limits.

Effects of Limits on Trading Systems

When developing a trading system, especially one that trades on the close, it is necessary to test prices for limit moves. This is particularly difficult because those limits have changed many times during the years. Using today's soybean limit of 70¢ would not allow the program to recognize any limit moves before 2011.

The easiest way to recognize a *locked-limit day* is that the open, high, low, and close are all the same. No trading occurred on that day. However, most days trade up to the
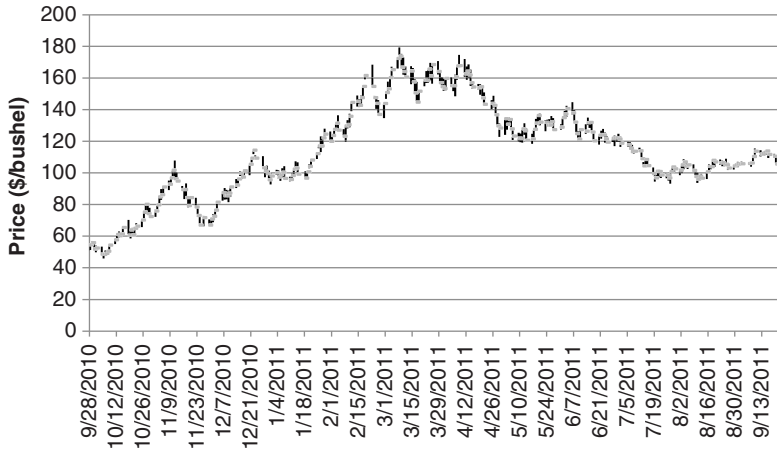
FIGURE 22.17 Cotton futures prices during the recent period of stress.

limit then stop, so that the high and close are the same, or the low and close, on a limit-down day. On those days you can test the change from the previous day. If that change is some even amount, say 20¢ or 40¢, there is a good chance is was a limit move. Allowing a trend system to change from a long to a short position on a limit day may produce a large gain when there may have been a large loss.

Most recently, cotton has been the center of extreme moves. Floods in Pakistan two years in a row, political disruption in Egypt, and drought in the southern United States have put pressure on prices, as shown in Figure 22.17. During February 2011 there are a number of days with gap openings representing limit moves. Figure 22.18 shows the closing price changes for the same period. Without knowing the daily limits, which can vary under certain circumstances, we can see that there are a large number of moves at even
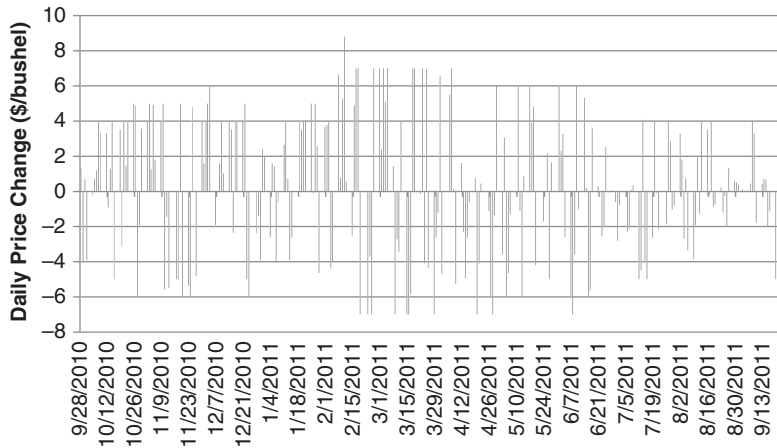
FIGURE 22.18 Cotton futures daily price changes.

number, $4, $5, $6, and especially $7 (the maximum limit), with one exception above $8 (a data error?). Given the frequency of these limit moves, ignoring them in an historic test of a system would give highly incorrect results.

Note that rules about limit moves can be different for each market and each exchange, and they have changed from time to time. In some cases, limits do not apply to the front month or the second month, so it is important to check the rules before trading.

Similar to Price Shocks

Limit moves are always the result of price shocks. If there is a comparable cash market open, then it will reflect the correct market price. In most cases, cash prices move quickly to the extreme, then start the slow, prolonged reversal. If you cannot liquidate a futures position because of a limit move, using the cash market to hedge the risk is not effective. The cash market would have already jumped to the price considered fully discounted by the traders. Buying in the cash market is more likely to add to your loss. Being long both futures and cash will simply double your risk. If there were no limits, then the full loss would have been seen on the day of the price shock. With limits, there is a chance that the some easing of the problem will occur and losses will be smaller by the time you exit the trade.

SILVER AND NASDAQ—TOO GOOD TO BE TRUE

This story of silver trading could have applied to the NASDAQ or S&P index during the late 1990s, or perhaps gold in 2011. The word "silver" could be replaced by "Intel" or by "Apple." There is no doubt that the same circumstances will reappear from time to time.

In 1974, public awareness of inflation caused an overwhelming interest in all forms of hedging, with large numbers of naive investors purchasing silver and gold as currency protection. Many sophisticated investors turned to the futures markets, which offered leverage on their purchases.

One investment system that was sold at that time was more of a leveraged substitute for the purchase of bullion than a trading strategy. At the time it was published, it had *always* worked—the sponsor of the system stood behind it with his reputation. The rules of the system were:

1. Trade silver futures because of their intrinsic value, historic performance, potential, and fundamental demand with short supply.
2. Use the futures contract between three and seven months from delivery to combine the advantages of liquidity and duration.
3. Always buy, never sell, because it is always successful.
4. Buy whenever you like. Although any sophisticated method can be used, it won't matter in the long run—any guess will do just as well. Follow the same method for adding to positions or reentering after closing out a trade.

5. Close out the position when there is a profit, not before.

6. Meet all margin calls—don't let the market beat you.

7. Invest $5,000 per contract (5,000 troy ounces). This will allow a $1 per ounce adverse move (silver was at $4.50 per ounce).

8. Whenever you need reinforcement, reread the reasoning behind this system.

9. Do not let anyone or anything interfere with following this system.

How did the investors do? It depended on when they started. Entering in the first half of 1975, shortly after the system was released, showed individual contract returns ranging from profits of 51¢ to losses of 18¢ per ounce. A drop in silver prices then produced losses of up to $1.02 per ounce for the next 9 months (slightly over 100%), followed by a 2-month rally, and losses again. This system would have worked from 1979 to February 1980 and then lost for the next 6 years. An investor following these rules, entering at a chance place, would have had more than a 10:1 likelihood of losing. Any profit made would have been given back if the investor persisted in following the method. Of course, had you held silver into 2011, you would have been profitable even after entering at the worse price—unless you consider inflation. Then prices again dropped by 25%.

The same results would have followed trading the NASDAQ. Buying anytime during the 1990s would have produced profits. No matter what trend method, indicators, filters, or dart board you use—if you buy and prices go higher, you always make money. If a trader continued to buy from 2000 through 2003, he would have faced the same total loss as the silver trader. The only true test of a strategy is the ability to make money in different types of markets.

In trying to understand how a system such as this can have a chance of reaching the public, remember the time in which it was introduced. With prices increasing drastically, food shortages, and publicity over the devaluing U.S. dollar, the rationale of the system seemed justifiable. The same arguments seem to apply to gold in 2011. There was serious talk about *leveraged inflation funds* using futures. We saw the same euphoria on the part of inexperienced traders who bought NASDAQ in the late 1990s. They had no expectation of losses. These systems fill an immediate need, without regard to long-term consequences. There is no doubt that, if high inflation or another major bull market returns, systems just like this one will reappear right near the top.

SIMILARITY OF SYSTEMATIC TRADING SIGNALS

A primary concern of both the government regulatory agency (CFTC) and individual traders is the similarity of trading signals generated by computer-based systems that are used to manage large positions. Throughout this book, we have noted that the market is the driver of trends, and that the specific trend-following method is less important. These systems may appear different in their rules and parameters, but they

could be essentially quite similar—for any of them to be profitable, they need to extract those profits from the same price moves, even though they may enter and exit as somewhat different points.

At the end of Chapter 3 there was discussion of how the S&P index has affected the price patterns of those stocks within the index. Traders enter orders to buy and sell the S&P based on S&P price support and resistance, or on trends applied to the S&P index. That, in turn, causes an arbitrage between the index and the individual stocks comprising the index; therefore, those stocks stop rising at the S&P resistance levels, even though that price may not align with the stock's own resistance point.

If systematic trend-following systems with different calculation periods are not highly correlated, then it is possible to reduce portfolio risk by trading a variety of different techniques and time periods; however, the reality may be different from the theory. Lukac, Brorsen, and Irvin[6] performed a study that compared 12 popular trading techniques (mostly trend-following) over 12 varied futures markets for the years 1978–1984. Each system was optimized, using 3 years of data, and the best parameter values were used to trade the next year. The systems selected were:

- Channel systems
  — Closing price channel (CHL)
  — MII price channel (MII)
  — L-S-O price channel (LSO)
- Momentum/oscillators
  — Directional Indicator (DRI)
  — Directional Movement (DRM)
  — Range quotient (RNQ)
  — Reference deviation (REF)
- Moving averages
  — Moving average with percent band (MAB)
  — Dual moving average (DMC)
- Systems with trailing stops
  — Parabolic Time/Price System (PAR)
  — Alexander's Filter Rule (ALX)
  — Combined Directional Movement and Parabolic Time/Price System (DRP)

The study used three measures to test system similarity:

1. The percentage of the time that systems are on the same side of the market (long or short).

[6] Louis P. Lukac, B. Wade Brorsen, and Scott H. Irwin, *Similarity of Computer Guided Technical Trading Systems* (CSFM-124, Working Paper Series, Columbia Futures Center, Columbia University Business School, New York, March 1986). Also by the same authors, "Do Similar Signals from Trading Systems Move Prices?" *Futures* (November 1987).

TABLE 22.6   Correlations of Aggregate Monthly Returns from 12 Systems

System	Trading System										
	CHL	PAR	DRM	RNQ	DRP	MII	LSO	REF	DMC	DRI	MAB
PAR	0.57										
DRM	0.72	0.61									
RNQ	0.70	0.41	0.70								
DRP	0.65	0.81	0.79	0.57							
MII	0.75	0.54	0.67	0.73	0.67						
LSO	0.59	0.43	0.53	0.68	0.54	0.70					
REF	0.55	0.37	0.57	0.66	0.52	0.54	0.60				
DMC	0.72	0.41	0.68	0.78	0.55	0.74	0.58	0.57			
DRI	0.71	0.42	0.69	0.77	0.55	0.70	0.59	0.66	0.64		
MAB	0.72	0.55	0.75	0.74	0.69	0.69	0.63	0.60	0.78	0.72	
ALX	0.58	0.55	0.62	0.55	0.57	0.57	0.58	0.51	0.52	0.56	0.57

All coefficients significant at the 0.01% level with the exception of REF and PAR coefficient, which is significant at the 0.05% level.

Source: Reprinted with permission of *Futures* magazine (250 S. Wacker Drive, #1150, Chicago, IL 60606, November 1987).

2. The percentage of buy or sell signals that occurred on the same day or within a few days of one another.

3. The correlation of aggregate monthly portfolio returns.

The results show a significant positive correlation in the system profitability (Table 22.6). However, there is no pattern that shows that one particular type of system is notably more correlated than others. The Parabolic and Directional Parabolic systems are most similar because one is based on the other. In Table 22.7, the percentage of trades that occur on the same day is very low. A more informative comparison is seen in Table 22.8, which tallies the percentage of days that each system held the same position.

The study concludes that computer-based systems trade on the same day significantly more often than would randomly be expected, but the actual percentage of trades that occur on the same day is small. These systems have the potential to move market prices. One must be concerned, as the assets under management of macrotrend trading programs get larger, that they compete with each other for both entries and exits, which now may span up to three trading days.

Comparison of Moving Average Systems

Using U.S. 30-year Treasury bonds as an example of a trending market, apply a simple moving average system to continuous futures prices from 2000 through mid-2011. The net profits of 6 calculation periods, from 10 to 60 days, are shown in Figure 22.19. Results seem to vary considerably.

TABLE 22.7   Percentage of Trades That Occur on the Same Day

					Trading System					
System	CHL	PAR	DRM	RNQ	DRP	MII	LSO	REF	DMC	DRI	MAB
PAR	19*										
DRM	25**	21**									
RNQ	22**	9	15								
DRP	30**	93**	48**	10							
MII	20*	15	10	28**	17						
LSO	22**	13	17	27**	18	23**					
REF	12	7	7	11	9	12	14				
DMC	16	6	8	28**	6	12	18	19*			
DRI	18	9	14	38**	10	32**	22**	15	25**		
MAB	28**	18	23**	23**	27**	18	25**	11	19*	23**	
ALX	12	20*	12	19*	20*	9	17	10	10	16	17

Percentage of trades that occur on the same day by each pair of systems. Significance assuming a binomial distribution with *denoting 95% confidence limits and ** denoting 99% confidence limits.

Source: Reprinted with permission of *Futures* magazine (250 S. Wacker Drive, #1150, Chicago, IL 60606, November 1987).

TABLE 22.8   Percentage of Trading Days Systems Hold the Same Positions

					Trading System					
System	CHL	PAR	DRM	RNQ	DRP	MII	LSO	REF	DMC	DRI	MAB
PAR	70										
DRM	82	76									
RNQ	75	61	73								
DRP	68	83	75	59							
MII	82	69	80	75	65						
LSO	73	58	71	73	57	73					
REF	72	57	69	69	55	72	69				
DMC	81	62	77	79	61	79	76	78			
DRI	70	56	68	72	54	71	69	67	73		
MAB	65	54	66	62	53	63	61	57	65	59	
ALX	72	64	71	68	59	73	65	67	73	63	58

Percentage of the trading days each pair of systems is on the same side of the market (long or short) at the same time. All coefficients are significant, assuming a binomial distribution with 99% confidence limits.

Source: Reprinted with permission of *Futures* magazine (250 S. Wacker Drive, #1150, Chicago, IL 60606, November 1987).

If we take the daily changes in PL for each of the six systems and find the correlations (see Table 22.9), we can see a clear pattern. Looking down in column 1, the 10-day calculation period, the correlations decline as the difference in the calculation periods increases. They also drop faster when the percentage difference in the calculation days is larger. Then the correlation difference between the 10- and 20-day trend is larger than the difference between the 50- and 60-day trends. This supports the statement in Chapter 21
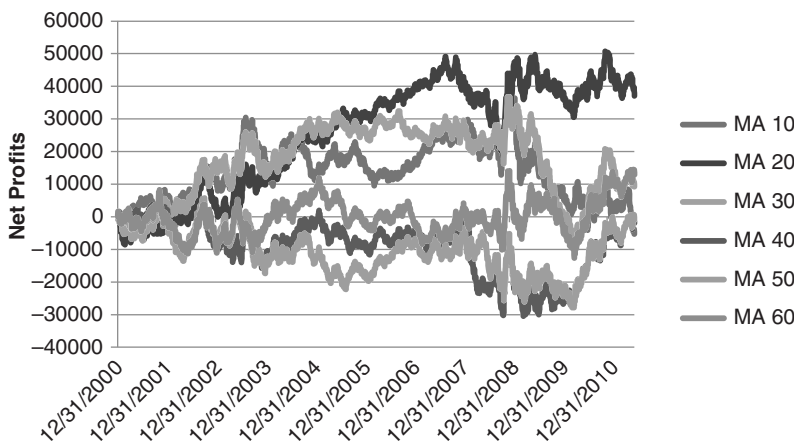
FIGURE 22.19 Net PL for six moving average systems, U.S. 30-year bond continuous futures, from 2000 through mid-2011.

that the calculation periods used for testing should not be in equal increments because they favor the slower periods. Also notice that the correlation between any two calculation periods will always be positive because they are extracting profits from the same trends. This was even more pronounced in the Lukac study over a period when the market had stronger trends.

Similarity of Long-Term Performance

The similarity of signals also implies the similarity of performance. In Chapter 21, it was shown that trend-following systems of about the same calculation periods had similar performance over time. The conclusion was that all of the methods identified the same trend, some sooner and some later, but the effect was that they all made money or lost money, to different degrees, during those trends. It was the market, not the strategy, that was most important.

TABLE 22.9 Correlations of PL Changes for Moving Average Systems of Different Calculation Periods, U.S. Bond Futures, 2000 through mid-2011

	MA 10	MA 20	MA 30	MA 40	MA 50	MA 60
MA 10	1					
MA 20	0.511	1				
MA 30	0.428	0.678	1			
MA 40	0.365	0.555	0.712	1		
MA 50	0.315	0.478	0.593	0.711	1	
MA 60	0.265	0.382	0.458	0.579	0.747	1

A recent paper by ISAM[7] attempts to show that all trend systems converge on the same performance over time. By defining three systems, but focusing on the *N*-day breakout, they show that various combinations of calculation periods and asset allocations cause variation in the year-to-year returns but total returns converge to similar results over time. Does this mean that it does not matter which parameters we choose, or how we allocate funds within a portfolio, because it will all come out the same in the end? Or, does it mean that, if we traded all combinations of parameters and allocations, that we would get an average result? Or, are the systems in the study not really representative of the set of trend-following systems that we have been discussing?

High Correlation at the Best and Worst Times

Different systems and the same system with different calculation periods will hold the same trade for varying periods of time. Typically, when they are in the same trade, prices are moving in a profitable direction. That is easy to understand. If a 10-day average was long and the 20-day average turned from short to long, then the 10-day would have continued to be profitable. In order to get the 60-day average to turn up, all of the trends of shorter periods are likely be profitable.

Once they are all in the same position, they also pose the greatest risk. The portfolio returns will increase and decrease by larger percentages each day, and, in the event of a price shock, all systems drop together. The process of controlling correlated risk, and overall risk, is discussed in the next two chapters.

---

[7] "Trend Following: The Myth of Return Dispersion," *International Standard Asset Management* (ISAM) (August 2011). See www.isam.com.

Risk Control

Every trading style has losing streaks that will ruin an investor who begins trading at the wrong time or is undercapitalized; therefore, the size of the position, the markets to trade, and when to increase or decrease leverage become important for financial survival. Some risks, called *systematic risk*, can be controlled or reduced while *market risk*, which can take the form of a price shock, can never be eliminated.

This chapter covers a broad range of topics related to risk, including individual trade risk, collective risk, leverage, and a continuation of the effects of price shocks and catastrophic risk. It is not possible to say that one is more important than another. In a specific situation, any one of the areas discussed may be the answer to preventing a substantial loss. The first part of this chapter discusses capitalization and shows why many traders are successful for months and then lose everything in only a few days. It explains the choices in dynamic leveraging and offers alternatives of less risk. Profit-taking and stops, the two most common ways to control risk, are shown to apply to specific types of trading, rather than used as a generalized solution. The last section tries to distinguish between a system that is temporarily performing properly and when it has failed to live up to expectations of it.

MISTAKING LUCK FOR SKILL

Embracing risk is more important than measuring risk. Understanding and accepting each day's losses and series of losses is essential for survival. Risk cannot be eliminated, as much as traders try to make each trade a profit or engineer systems to minimize losses. You can move the profits and losses around, but you cannot eliminate them or make them so small that they are meaningless. And measuring risk as a 10% chance of losing 20% of your equity is not as sobering as actually losing 20%. It is even devastating when you thought you could only lose 5%.

Daily market volatility gives a good approximation of short-term risk, and volatility of the actual profits and losses of a trading method provides a way to estimate how

much risk must be absorbed to achieve the expected profits. Price shocks are the fly in the ointment. They add an unpredictable aspect to risk, coming at unexpected times and cause losses that are larger than anything seen in the past. The previous chapter continued a discussion of some practical aspects of price shocks, with examples of how to manage both profits and losses caused by these large moves. This review only serves as a reminder of a very important problem.

A price shock represents the greatest risk for traders. In most cases, a price shock causes a loss, but when it is a windfall profit it can be disarming.[1] A price shock, such as 9/11, cannot be anticipated, and any profit that might have been gained is the result of luck, not skill. Traders must realize that the profit could have been a loss; they should assess how their account would have survived a drop of the same magnitude. Mistaking luck for skill ignores the risk and exposes the trader to future risk for which he or she is now unprepared. Long-Term Capital Management, a home for some of the smartest financial minds, believed they could engineer the risk out of each trade by explaining away all of the previous price shocks, then failed to survive the next price shock. Without including these larger, infrequent losses in the performance profile, trading accounts are likely to be undercapitalized.

The only way to reduce the possibility of loss from a price shock is to be out of the market as much as possible. This translates into two practical trading rules:

1. *Do not hold a position longer than necessary.* Choose a strategy that has good entry timing and has many opportunities for exiting the trade. It is best to be in the market when prices are moving and out when volatility drops and prices are no longer active. A trading method that is in the market only 40% of the time has a 60% chance of avoiding a price shock. A strategy that is always in the market will be exposed to all price shocks.

2. *Try to earn as much as possible while investing as little as possible.* The less you have invested, the less you have at risk. Futures and options offer an opportunity to leverage price moves by as much as 20:1; however, that works against you when a price shock causes a large loss. A fully leveraged trading program that has a theoretical return of 200% per annum will also be a total loss after an adverse price shock. A deleveraged program may return only 20% but is 10 times more likely to survive a price shock, a much better long-term alternative. A small reduction in returns for a larger reduction in exposure is a good trade-off.

RISK AVERSION

Traders react differently to risk, and most are risk-averse. A primary reason for the popularity of trend systems is that they provide some degree of *conservation of capital*. They

---

[1] A book entirely devoted to price shocks and the "black swan" is Nassim Nicholas Taleb, *Fooled by Randomness* (Texere, 2001). Taleb has popularized the phrase "mistaking luck for skill."
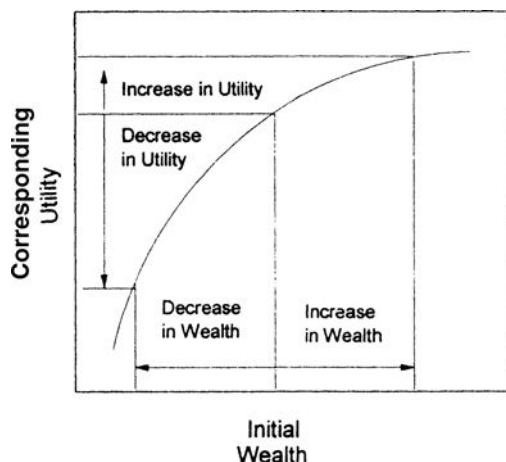
FIGURE 23.1 Changes in utility versus changes in wealth.

Source: Peter Bernstein, *The Portable MBA in Investment* (New York: John Wiley & Sons, 1995), 38. © 1995 Peter Bernstein. Reprinted by permission of John Wiley & Sons, Inc.

take small losses when prices move the wrong way, and hold a trade for much longer when it shows a profit. By taking small losses, the investor can wait until a sustained trend comes along. This gives the trending system its familiar profile of a large number of small losses offset by a few large profits. But each trader has her personal tolerance level for risk and not all of them find conservation of capital appealing. Some are risk seeking.

Daniel Bernoulli, a famous mathematician, proposed a theory of utility in 1738 that distinguished between price and value, where price is the same for everyone but value (utility) depends on the personal circumstances of the individual making the estimate.[2] Bernoulli's approach defined a concept of *diminishing marginal utility*, shown in Figure 23.1, which indicates that as wealth becomes greater, the preference for more wealth diminishes (unlike the concept of power). In the lower left part of the chart, where the investor has a small net worth, the likelihood of accepting risk is much higher, although the magnitude of the risk is still small. When risk becomes greater, even proportional to reward, all investors become cautious. Bernoulli's graph shows the curve beginning at zero and moving up and to the right in a perfect quarter-circle, ending horizontally where risk is no longer attractive. This implies that people are risk-averse. Most people are not interested in an even chance of gaining or losing an equal amount. Other theories that have been proposed are that the market maximizes the amount of money lost and the market maximizes the number of losing participants. All of these concepts appear to be true and are significant in developing an understanding of how the market functions.

---

[2]Peter L. Bernstein, ed., *The Portable MBA in Investment* (New York: John Wiley & Sons, 1995), 37.

Naturally, there are exceptions to most rules. People will spend small amounts to have a very small chance of winning a very large amount, as with various state lotteries. Young people with a small amount of money are willing to take greater risks, rationalizing that a small amount of savings is not significant for their future.

Risk Preference

The Bernoulli theory of utility recognizes that each investor has his or her unique objectives and attitudes toward risk. Some investors would like to keep risks low and returns steady; more speculative investors would risk a sizable amount of their capital for a chance at the big move. This trait is called the investor's *risk preference*. The *risk preference*, *P*, or *utility* of an investor for a specific venture (in this case a trade) can be found by adding the expected value of the investor's utilities or preferences, p_i, for the various outcomes of that event,

$$P = w_1 p_1 + w_2 p_2 + \cdots + w_n p_n$$

where $\sum w_i = 1$
$n = $ possible outcomes

The weighting factors may be the results of personal bias or may be the calculated probabilities of each outcome. For example, consider a gold trade that has a likely profit of $4,000 and a risk of $1,500. For convenience, adjust the reward values by dividing by 1,000. If the probability of success is 60%, the total utility of the trade is:

$$P(\text{trade}) = 0.60 \times 4 + 0.40 \times (-1.5) = 1.8$$

If the probability of success were increased, the utility *P* would increase linearly. But investors do not feel the same about different rewards. Given a scale of 0 to 100 (negative to positive reaction), an investor may rank a 60% chance of a $4,000 profit and a 40% chance of a $1,500 loss both as a 65. If the reward is increased to $8,000 while the risk remains at $1,500, the investor might only raise the preference of the trade to 80, although the utility would be 4.2, more than twice as large.

The various curves drawn through the computed utilities represent the risk preference of the individual. Figure 23.2 shows these curves progressing from extreme risk aversion to extreme risk seeking. As the risk increases in (1), the trader is risk-averse and less likely to participate in the trade; in (3), there is equal chance of taking the trade at all risk levels; and in (5), the trader is risk seeking and is more likely to enter a trade that has higher risk.

The Efficient Frontier

There are a large number of performance profiles that result from combining trading strategies, a diverse selection of stocks, futures, ETF, and funds, and the amount of leverage used. Each of these investments can be described in terms of its annualized rate
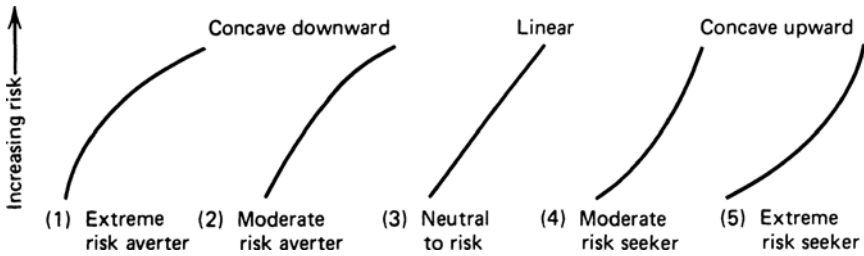
FIGURE 23.2 Investor utility curves.

Source: R. J. Teweles, C. V. Harlow, and H. L. Stone, *The Commodity Futures Game: Who Wins, Who Loses, Why?* (New York: McGraw-Hill, 1974), 133. © 1974. Reprinted with permission of The McGraw-Hill Companies.

of return and its annualized risk. Risk is most often the standard deviation of the annual returns; however, other measurements will be discussed in the next sections.

The *efficient frontier* is a way of visualizing which choices would be made by any rational investor. Suppose there were 8 different investment programs, and all of them reported their annualized rate of return and the annualized risk. These could then be plotted in Figure 23.3. Let us also assume that these investments represented a good sample of all risk and return combinations.

A *rational investor* is one who wants the highest return for the lowest risk. Therefore, given a choice between investment *D* and investment *C*, we would choose *C* because it had the same risk but a substantially higher return. Similarly, if we were given a choice between investment *D* and investment *A*, we would choose investment *A* because it has the same return with a much lower risk. Given a choice of any two investments, displayed on a chart of risk and return, a rational investor will always choose the one that is higher and to the left.

The efficient frontier is a curve plotted through those investment alternatives that have the highest returns for a given risk, or the lowest risk for a given return. An investor
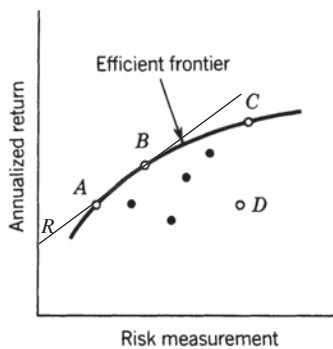
FIGURE 23.3 The efficient frontier. The line drawn from *R*, the risk-free return, touches the efficient frontier at the point of best return to risk ratio, *B*.

seeking greater returns with a higher tolerance for risk will choose an investment further to the right on the curve; one with a lower risk tolerance must select a program farther to the left and lower, with less return.

The optimal point on the efficient frontier is where the return-to-risk ratio is a maximum. This is also the point where a line, drawn from the risk-free return on the left scale upwards to the right, is tangent to the curve. Nevertheless, not all investors choose this point because the absolute level of risk may be too high, or the absolute returns too low. Investors' risk preferences are personal and unique. Finding the optimal portfolio selection is discussed in detail in Chapter 24.

Common Sense Management of Risk

Risk control is as much an issue of common sense as it is complex rules and mathematics. Most of this chapter shows how various rules, measurements, and leveraging techniques can reduce risk for both discretionary and algorithmic trading; however, successful traders have applied common sense, without complex formulas, for a long time. Some of these successful principles are:

1. *Only risk a small amount of total capital on any one trade.* The total amount risked should allow you to comfortably survive a number of losses in a row. No trade should ever risk more than 5% of the invested capital.

2. *Know your exit conditions in advance.* There should be a clear exit criterion for every trade, even if the exact loss cannot be known in advance.

3. *Large profits mean large risk.* If the average profits or average losses are too large relative to the investment, then smaller positions should be taken.

4. *Exit a trade quickly.* Exit a trade as soon as you recognize that it has gone wrong. Don't try to manage the loss. Many floor traders believe that the smartest trader is the first one out.

5. *Don't meet margin calls.* Experienced traders believe that a margin call is an objective statement of a trade that's gone wrong, or a system that is not meeting expectations. It is a time to review trading performance rather than invest more.

6. *Liquidate your worst position first when lightening up.* The profitable trades have proved that they are trending or performing properly; the losing ones have proved they are not. Stay with the good positions and liquidate the worst.

7. *Be consistent with your trading philosophy.* If you are a trend follower, then keep losses small and let profits run. You cannot be a trend follower by taking profits whenever they occur.

8. *Be sure the trading profile is compatible with your risk preference.* You cannot follow a strategy that takes risks that are uncomfortably large. Be sure that profile is agreeable to you.

9. *Plan for contingencies.* Nothing ever goes as planned, and you must be prepared for infrequent but important exceptions, such as a price shock. Do not be undercapitalized.

LIQUIDITY

The ability to implement any trading program is directly related to the liquidity of the market. Even though liquidity can have more than one meaning, lack of liquidity has only one trading result—poor execution. A *liquid* market does not necessarily mean good fills, but an *illiquid* market assures bad ones. Two types of illiquidity most often encountered in the markets are:

1. *Fast markets*, in which there are mostly buyers or sellers, and few traders willing to take the other side of the position.
2. *Inactive markets*, usually the smaller cap stocks, deferred futures contracts, and many ETFs, where there is less commercial or institutional interest.

The fast market is the result of developing news or an economic report where everyone has the same interpretation of the direction that prices should move. In futures this may be combined with a supply/demand imbalance or a perceived imbalance. Once a large number of buy orders flood the market, specialists, market makers, and locals all stand aside (or widen the bid-asked spread and offer very small quantity) rather than be run over. The few arbitrageurs, hedgers, and risk-seeking traders willing to take the opposite position do little to offset the vast number of orders that continue to push prices in one direction. In the inactive market, a large premium must be paid to attract another trader to take the opposite side. This only succeeds when the bid or offered price clearly appears to give the other trader a guaranteed short-term profit.

An illiquid market causes poor executions, which is the largest part of transaction costs and is sometimes too large to overcome. Ginter and Richie[3] have described this as a function of the order size, volume, and the speed of the market, in a single formula that applies to the futures markets, but could be adapted to individual equities:

$$C = K \times \frac{Q \times V}{L_c \times L_o}$$

where C = cost of execution due to liquidity
        Q = size of the order entered
        V = volatility of the market
      L_c = volume of the market (all contract months for futures or possibly the
            combined volume of highly correlated equities)
      L_o = volume of the specific delivery month or individual equity
      K = constant factor

---

[3] G. Ginter and J. Richie, "Data Errors and Price Distortions," in Perry J. Kaufman, ed., *Technical Analysis in Commodities* (New York: John Wiley & Sons, 1980).

The volatility V might be the daily price range, annualized volatility, or implied volatility. This provides a way to measure how much volume would move the market a specified number of points. The total volume (liquidity) L_c is important because it implies liquidity due to interdelivery spreading in futures. If applied to stocks, added liquidity may come from activity in a closely correlated industrial group. Volume also serves as a measurement of the general interest in the product or in the related stock group. Increased trading in other futures market delivery months, or in similar stocks, is a sign of growing interest.

The constant factor K varies according to the type of the execution order being placed and the current direction of the price with regard to that order. If prices are moving higher and a buy order is placed, K will be large; if prices are moving higher and a sell order is placed, K will be smaller. The investor can see that transaction costs, including commissions and execution slippage, will have considerably greater impact on systems that trade more often, have smaller average profits, and trade in the direction of the trend. Given a choice between systems of equal gross profits, the investor should favor the method of fewest trades in the most liquid markets.

MEASURING RETURN AND RISK

Risk and return are the two most important criteria for measuring performance and for determining that one trading method is better than another. Looking at only returns is not enough. However, the ways in which these and other measures can be combined to express the quality of performance varies considerably. For example, returns are often the annualized rate of return (see Chapter 2), but they can also be expressed as the slope of a straight line drawn through the daily or monthly returns. Risk has many more variations from the annualized standard deviation of returns to measurements of drawdown and time to recovery. The following section gives an overview of these techniques.

Risk Characteristics of a Trading Model

Before either risk or returns can be measured, it is necessary to have a strategy that has produced historic results, whether real or theoretical. The use of actual trading results is always better because it will show larger equity swings than hypothetical performance, regardless of how conscientiously you have performed the development and testing. Hypothetical results are always optimistic. In some cases, techniques such as analyzing trends in returns do not work with theoretical results but can be a great benefit in real trading.

Suppose there is a choice of systems to trade, and the daily equity of each is available. If only one can be selected, which should it be? There are some characteristics of performance that are normally considered better:

1. Larger profits rather than smaller profits because it leaves more room for execution error.
2. Upwards equity surges (profits) over downward surges (losses).

3. A shorter time-to-recovery rather than prolonged drawdowns.

4. Less time in the market because it avoids price shocks.

Results are always relative. If your choice is between a buy-and-hold portfolio in stocks with a return of 8% and risk of 11%, compared to a simple moving average system of S&P futures with 12% returns and 13% risk, then the moving average system is clearly better, although far from being a money machine. To find the best choice always include a benchmark. Most often this is the result of a theoretical portfolio of 60% stocks (using the S&P) and 40% Treasury bonds (using the Lehman Brothers Treasury Index).

The Sharpe Ratio and Information Ratio

The classic measurement of performance is the *Sharpe ratio* (SR), expressed as

$$SR = \frac{E - I}{\sigma}$$

where E = the expected return (the annualized rate of return)
      I = the risk-free interest rate (usually the 3-month rate)
      σ = the standard deviation, or fluctuation, of the periodic returns (the annualized volatility)

Calculations for the annualized rate of return and annualized volatility are given in Chapter 2. For practical purposes, I is often omitted, with the simplified form called the *information ratio*, the most common performance statistic. The risk-free return, I, cannot be omitted when there is interest income on unused capital, or on some part of the account balance, that has been imbedded in the performance. For futures trading, interest can be earned daily on a large part of the market-to-market value of the account. For a comparison of different investments, it is necessary to specifically identify which returns include interest income. In the world of Commodity Trading Advisors (CTAs), there may be a performance hurdle, such as the yield on the 3-month Treasury bill, before any incentive is paid.

The Sharpe and information ratios satisfy the first universal criterion of system selection, that all else being equal (primarily risk), higher profits are better. It does not satisfy either of the other criteria illustrated in Figure 23.4 because it cannot distinguish between:

1. Consecutive small losses (System B) and alternating small losses (System A), shown in Figure 23.4a.

2. Large surges of profits and large losses, shown in Figure 23.4b.

Clearly, System *A* is best in both cases. Also note that performance of monthly returns is always better than results based on daily returns because it is inherently smoothed. It

(a)

(b)
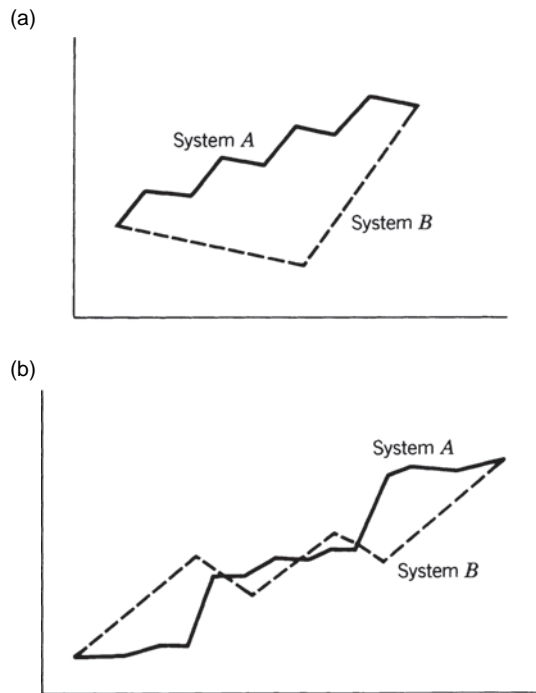
FIGURE 23.4 Two cases in which the Sharpe ratio falls short. (a) The order in which profits and losses occur. (b) Surges in profits versus evenly distributed losses.

is unlikely that the end-of-month net asset value (NAV) will be either the lowest or highest of the month, although highest is more likely than lowest.

Treynor Ratio

An additional measure of return to risk is the *Treynor ratio* (TR), the annualized return less the risk-free return, all divided by beta, the relative volatility of the current portfolio compared to a benchmark, usually the S&P

$$TR_t = \frac{AROR_t - I}{\beta_t}$$

Average Maximum Retracement

Schwager[4] has presented a comprehensive study of risk and return evaluation techniques. Although each method may emphasize a particular equity trait, he seems to favor

---

[4] Jack D. Schwager, *A Complete Guide to the Futures Markets* (New York: John Wiley & Sons, 1984).

the *Average Maximum Retracement* (AMR). This method finds the average equity draw-down, the difference between the current value in the account and the highest past value, by ignoring all days when equity was on new highs.

$$AMR_t = \frac{1}{n} \sum_{i=t-n+1}^{t} POS\left(MCE_i - TE_i\right)$$

where $\quad POS(X) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases}$

and $\quad MCE$ = closed-out equity (realized profits) on any trade entry date prior to i
$\qquad TE_i$ = total equity on day i
$\qquad n$ = total number of days of equity data or the calculation period

When $TE_i > MCE_i$, then equity is on a new high and all traders will have a profit on day i, regardless of when they began. Schwager suggests that a much simpler computation would use only the low equity day of each month; it would give a rough but good approximation.

Measuring only the drawdowns is a concept similar to *semivariance* and is valuable because the distribution of profitable returns compared to losing returns is not symmetric; therefore, the standard deviation does not express it correctly and may bias the results in favor of profitable days, giving you an understated picture of risk.

Largest Loss and the Calmar Ratio

Measurements such as Schwager's AMR, and even the basic standard deviation, are good for comparing the long-term performance of one system against another; however, they lack a certain reality of simply looking at the largest loss seen over the test period. John Sweeney calls this the *maximum adverse excursion*,[5] advocating that traders should minimize the size of their largest loss.

Consider the standard deviation of equity changes, showing that in any month there is a 68% chance that your returns will be between 15% and −5% (a mean of 5% and a standard deviation of 10%). There is only a 2.5% chance that you will lose more than 15% in one month (2 standard deviations, one sided); therefore, there is a 50% chance you will lose that 15% in one of the first 20 months (20 ÷ 2.5). Yet probability shows that you should lose more if you keep trading or less if you stop sooner.

The largest historic loss, called the *maximum drawdown*, is a practical alternative. It simply states that the trading program *did* lose 15% during one month of a 3-year test. While it is possible, and even likely, that the program will have a larger loss in the future,

---

[5] John Sweeney, "Maximum Adverse Excursions for Stops," *Technical Analysis of Stocks & Commodities* (April 1987).

you must be prepared for a 15% loss in a single month. The relationship of the drawdown to the returns are expressed as the *Calmar ratio*

$$Calmar\ ratio_t = \frac{AROR_t}{Max\ drawdown_t}$$

where *max drawdown* as of day t is the largest historic drawdown, peak to valley, from the beginning of the data to today.

Using the historic drawdown is practical but not without problems. In the previous chapter on testing, there were many cases where final system profile benefitted from unintentional overfitting. Even when parameters were chosen from a robust section of the test surface, they might all have looked good because the timing avoided a particularly bad price shock. Then the theoretical maximum drawdown might be understated. Even if not overfitted, then the future always brings larger profits and especially larger drawdowns.

Sortino Ratio

A variation on both the Sharpe ratio and semivariance, the *Sortino ratio* includes only the downside risk in the denominator and replaces the risk-free threshold with a minimum acceptable return (MAR) in the numerator. Then the Sortino ratio is calculated as

$$SR_t = \frac{AROR_t - MAR}{\sigma(PE_t - E_t)}$$

MAR may be the risk-free rate of return or any other threshold set by the investor. The denominator is the standard deviation of all the negative equity excursions, the difference between the peak equity (*PE*) and the current equity (*E*), not equal to zero, on each day during the calculation period.

Ulcer Index

Investors have increased anxiety as current returns drop farther below the highest returns previously achieved. This can be measured by the *Ulcer Index* (*UI*)[6] a form of semivariance, similar to Schwager's *Average Maximum Retracement*, that produces a statistical measure of relative declines on all days that were not new high returns.

$$UI_t = \sqrt{\frac{\sum_{i=t-n+1}^{t} D_i^2}{n}}$$

---

[6]Created by Peter Martin in 1987, it appeared in a book by Peter Martin and Byron McCann, *The Investor's Guide to Fidelity Funds* (New York: John Wiley & Sons, 1989).

where D = the difference between the highest equity as of day i and the actual equity
on day i

n = the number of days in the equity stream

If the equity on day i is a new high equity, then $D_i = 0$. As *UI* increases, investors are more anxious about performance.

Potential Risk

Risk is never a single value; it is a probability, and the standard deviation is the tool most often used for finding that relationship. Even if a drawdown has occurred that is larger than the previous maximum drawdown, there is still a small chance that another, even larger drawdown will occur. It is unreasonable to think that all future losses will be smaller than the maximum already experienced. This *potential for loss* can be expressed as a probability with two different measurements over n days:

1. *Probability of a drawdown* (DP). This is essentially the method used by Schwager and the Ulcer Index. Calculate the standard deviation of all daily drawdowns, D_i, measured from the most recent equity high to today's equity value. When today's equity is also a new high equity, the value used is zero.

$$DP = stdev(D_i)$$

where $i = 1$, n, and n is only those days where a drawdown existed, that is $D_i \neq 0$.

2. *Semivariance (SV)*. Calculate the linear regression of the equity stream or *NAV*, then find the standard deviation of all drawdowns, D_i, below the corresponding value of the straight line fit, p'_i. Semivariance will produce a smaller value than DP because the values on the straight line will be lower than the peak equity.

DP is a more conservative version of both Schwager and the Ulcer Index because it omits those days on which the equity made a new high. The standard deviation will be increased in proportion to the number of days omitted. With either measure, DP and SV represent the value of 1 standard deviation. Therefore, there is a 16% chance that there will be a drawdown greater than DP or SV over the next n days. There is only a 2.5% chance of a drawdown of twice that amount over the next n days.

Either of these measurements can be used to form a ratio of return to risk, called a *drawdown ratio* (DR):

$$DR_t = AROR_t/DP_t$$

or

$$DR_t = AROR_t/SV_t$$

where $AROR_t$ = the annualized return on investment on day t.

These ratios satisfy all three of the original criteria: higher profits are favored because the rate of return is in the numerator, the order of profits and losses will result in larger and smaller net equity drops, and large gains are not penalized because only the drawdowns are used to measure risk.

Conservative investors may want to include some additional simple considerations of potential risk. All else being equal, systems will have greater risk if:

1. They are tested with smaller samples.
2. They have few historic equity drops (also likely to be a poor sample).
3. They concentrate on fewer product groups (not well diversified).
4. They compound positions.

Annualized Volatility

The *annualized volatility* is mentioned separately because is it the most common way to express risk. It is simply the annualized standard deviation, previously given in Chapter 2 but used throughout this book.

$$V_t = \sigma(r_i) \times \sqrt{252}, \; i = t - n + 1, \; t$$

where r represents the returns for the past n days and 252 is the number of trading days in the year. Analysts simply call this *volatility*. Even though we know that returns are not symmetric, this remains the most common measurement of risk. In portfolio analysis, this will be used to determine the *target volatility*.

Value at Risk

Value at Risk (VaR) is used in most companies to assess whether the current market positions are likely to produce a loss that is unacceptably large over the next few days. It is different from volatility given in the previous section because it attempts to *anticipate risk*, even though it is done using historic risk. This is done by finding the probability of loss of the current market positions based on past market movement. For example, Bank Two holds large short positions in fixed income futures to protect loan commitments at current levels. They also hold a wide variety of foreign currencies, of which 50% are hedged with forward contracts. If interest rates drop or the U.S. dollar strengthens, they could show large losses in their futures positions. According to bank policy, the loss allowed in a single day must be kept to under 0.5% of cash value of these commitments. If there is a potential risk exceeding that amount, the futures position must be immediately reduced.

VaR is a combination of cross-correlations between markets for which there is exposure, the position sizes, the volatility of those markets, the projected time period over which the risk will be forecast, and a confidence interval to determine the risk tolerance. Using the return series for the interest rate and FX exposure held by Bank Two, the *variance-covariance* solution for VaR can be calculated as[7]

[7]Jacques Longerstaey, *RiskMetrics™ Technical Document*, 4th ed. (Morgan Guaranty Trust Company, riskmetrics@jpmorgan.com).

$$VaR = \sqrt{\sigma_{IR}^2 + \sigma_{FX}^2 + (2 \times \rho_{IR,FX} \times \sigma_{IR} \times \sigma_{FX})}$$

where σ = the standard deviation of the return series for the individual markets

        p = the cross-correlation between the two return series

In this example, the return series reflects the position size and volatility of the exposure. This formula implies a VaR projected period of 1 day.

For example, Bank Two holds 1 million in EURUSD, which have a daily standard deviation of 0.565%. If the current EURUSD rate is 1.25, then the market value of the position is USD 1.25 million. It also holds a EUR 1 million position in 10-year Eurobunds, which show a standard deviation of 0.605% and a market value of USD 1.25 million. Using 1.65 standard deviations to represent the 95% confidence level, and making the assumption that r_i/σ_I is normally distributed, and the EURUSD exchange rate should not drop more than $0.565 \times 1.65 = 0.932\%$ (less than 1%) on a single day, 95% of the time. The Eurobunds should drop less than $0.605 \times 1.65 = 0.998\%$ on a single day. We could then expect the approximate risk of the FX position to be USD 1.25 million $\times$ 0.932% = USD 11,650 and the risk of the Eurobund position to be $1.25 million $\times$ 0.998% = USD 12,475 over the next 24 hours.

The total risk of the FX and Eurobund positions is not the sum of the two risks, because correlations between the two markets may show that price movement of one market is partially offset by opposite movement in the other market. In this example, the correlation of the returns of two markets is –0.27, showing that there is a noticeable offsetting effect. Applying these values to the previous formula gives

$$VaR_t = \sqrt{12475^2 + 11650^2 + (2 \times -0.27 \times 12475 \times 11650)}$$

$$VaR_t = \text{USD } 17,068$$

The value at risk for the next 24 hours is far less than the sum of the two individual market risks because their price movements tend to offset each other.

A more general calculation of VaR uses separate values for the underlying prices, the position size, the time period, and the confidence interval. For three assets, this is

$$VaR_i = CI \times \sqrt{P}$$

$$\times \sqrt{w_1^2 x_1^2 \sigma_1^2 + w_2^2 x_2^2 \sigma_2^2 + w_3^2 x_3^2 \sigma_3^2 + 2w_1 w_2 x_1 x_2 \sigma_1 \sigma_2 r_{12} + 2w_1 w_3 x_1 x_3 \sigma_1 \sigma_3 r_{13} + 2w_2 w_3 x_2 x_3 \sigma_2 \sigma_3 r_{23}}$$

where CI = the confidence interval (expressed as 1.65 standard deviations for a 5% probability

       P = the VaR calculation period (5 is a projected period of 5 days)

       w_i = the weighting factor, or relative position size

       x_t = the current value or price of the asset on the current day t

       σ_i = the annualized volatility, shown below

       r_{ij} = the cross-correlation between market i and market j

Cross-correlations can be found using the spreadsheet function *correl* or can be calculated directly from the formula in Chapter 6. The volatility, σ_i, is calculated using the prices, x_i, over n days, as the standard deviation of the price changes, annualized:

$$\sigma_i = \sqrt{252} \times stdev(d_i)$$

$$d_i = \frac{x_t - x_{t-1}}{x_{t-1}} \text{ over the past } n \text{ days}$$

The square root of 252 is used to annualize the daily values.

In the practical implementation of risk control, VaR would be calculated before the close of trading and, if its value exceeds the threshold amount, action can be taken to reduce the exposure on the same day.

Generalized VaR Calculation

The probability of loss at the end of a forecast period, P_E, is calculated as the difference between the cumulative percentage loss over that period and the cumulative expected return, dividing this value by the cumulative standard deviation, and applying the standard deviation of a normal distribution to give a probability estimate.[8]

$$P_E = N\left[\frac{\ln(1+L) - \mu T}{\sigma\sqrt{T}}\right]$$

where $N[\ ]$ = the cumulative normal distribution function

L = the cumulative percentage loss in periodic units

μ = the annualized expected return in continuous units

T = the number of years in the horizon, where 1 day is $\frac{1}{252}$

σ = the standard deviation of continuous units

In this process, the compounding of periodic returns results in a lognormal distribution, while the continuous returns and standard deviation are normally distributed. To calculate VaR, this calculation is changed to specify the probability, Z, and solve for the size of the threshold loss.

$$VaR = -(e^{\mu T - Z\sigma\sqrt{T}} - 1)W$$

where e = the base of the natural logarithm (2.7128)

Z = the chosen probability deviate (for example, 5% is 1.65)

W = the initial value of the asset

The result of this calculation gives the probability of loss at the end of the period, T, and does not reflect any losses that might have occurred during the period (if T is greater

---

[8] Mark Kritzman and Don Rich, "The Mismeasurement of Risk," *Financial Analysts Journal* (May/June 2002). This entire section uses material from the Kritzman and Rich article.

than 1 day). Losses during the period will always be at least as large as the loss at the end of the period.

A Simpler Value at Risk

There are three calculation methods associated with Value at Risk: variance-covariance, historical, and Monte Carlo. The section above discussed the classic variance-covariance approach, but most analysts will find the historical method easier to implement and intuitively simpler. In essence, to calculate the VaR for today, take today's market positions and apply them to the data from the past, say, 1 year (252 days). Specifically,[9]

1. Given n markets or assets, each with positions (contracts or shares) $w_1, w_2, \ldots, w_n$.
2. Calculate the returns for the past k days ($k = 252$) as $r_i = p_i/p_{i-1} - 1$, $i = t - k + 1, t$.
3. Calculate the total returns for day i as $R_i = w_1 r_{1i} + w_2 r_{2i} + \ldots + w_n r_{2i}$. Include the forex and conversion factors (for futures) if applicable so that the total returns are all in USD, euros, or any single currency.
4. Find the average, $\overline{R}$, and standard deviation, s, of R over the k days.
5. Given a confidence level specified in standard deviations (e.g., 99% = 2.33), the VaR for the current day, t, is $\overline{R} - 2.33s$. VaR is always expressed as a negative number.

Price Shocks and Realistic Distributions

Value at risk is a good indication of risk under most market conditions, but is known to be a poor measurement during extreme, highly correlated price moves across market sectors. It has been said that VaR works *except* when you need it the most—when the risk is greatest. A *stress test* is used to find out how VaR performs under extreme cases. This isolates specific market periods, such as price shocks (discussed in the previous chapter), or can use simulated data in which all prices reverse direction and make a 2-to-3 standard deviation move with very high correlation, as in the subprime collapse in late 2008.

The assumption of a normal distribution in either returns or price movement is not realistic, an opinion held by Benoit Mandelbrot, among others, who maintained that using the standard deviation of equity changes underestimates the real risk of the financial markets.[10] Most return distributions are skewed and have a fat tail due to the application of conservation of capital and the unusually long, nonrandom distribution of runs. A log distribution could be substituted for a normal distribution, but this tends to be more extreme than necessary. The degree of fat tail present in the distribution can be measured using *kurtosis*,

$$k = \frac{\sum_{i}^{n} r_i^4}{n\sigma^4}$$

[10]B. B. Mandelbrot, "A Multifractional Walk Down Wall Street," *Scientific American* (February 1999).

where r_i = the return on day i
σ^2 = the variance
n = the number of days

A normal distribution has a kurtosis of 3, but the actual returns series might have a kurtosis as large as 25, indicating a significant fat tail. Although the non-normal distribution can be identified, there does not seem to be a simple solution for the form of the distribution that can be applied to VaR. When extreme distributions exist, the practical solution is to perform stress tests to recognize the potential risk during extreme cases and treat these as exceptions in the same way that crisis management might be used during a price shock. Advanced work with VaR focuses primarily on alternative distributions and unexpected, extreme risk.

Hindenburg Omen

The *Hindenburg Omen*[11] is a price pattern that attempts to forecast a stock market crash. It is named after the famous Hindenburg disaster of May 6, 1937, during which the German Zeppelin *Hindenburg* exploded and burned. The basis for the calculation is Norman Fosback's *High Low Logic Index* (HLLI), which is the lesser of the NYSE new highs or new lows divided by the number of NYSE issues traded, then smoothed using an exponential moving average. The premise is that, under normal conditions, a large number of stocks either make new highs or new lows, but not both at the same time. The existence of both indicates a specific case of potential volatility. Specifically, on a given day,

- The number of NYSE new 52-week highs and the number of new 52-week lows are greater than or equal to 2.8% of the total issue that advance or decline that day. Note that the percentage 2.8, was raised from a previous level of 2.5 as overall market volatility has increased.
- The NYSE index is greater than it was 50 trading days ago.
- The McClellan Oscillator is negative (see Chapter 12).
- The new 52-week highs cannot be more than twice the 52-week lows, although the new lows may exceed the new highs by any amount.
- The combination of events must occur on 3 consecutive days.

Once the signal has occurred, a stock market crash is expected within 30 trading days, but only when the McClellan Oscillator is negative.

Depending on your definition of a "crash," a drop between 5% and 50%, you could expect the Hindenburg Omen to be successful about 25% of the time. Most recently, these included August 12, 20, 24, 25, and 31, 2010.

---

[11]According to Wikipedia, the Hindenburg Omen originated with Jim Miekka, and the name was suggested by the late Kennedy Gammage.

Risk Characteristics of Systems

Every trading method has its unique risk and return pattern. The difference between trend and mean-reverting systems was discussed in the section "System Trade-Offs" in Chapter 22. In general, a trend system has many smaller losses and fewer larger profits. A mean-reverting system has the opposite; many smaller profits and fewer larger losses. This is similar to the choices available with gambling strategies.

Even among trending systems there are enormous differences in trading risk. A typical moving average system, which qualifies under *conservation of capital*, keeps individual trade losses small compared to individual profits, but may have a series of losses that when combined, add up to one very large loss. In addition, trend systems fail when they encounter a sideways period, and frequently reverse from long to short and back again until prices finally pick a direction. Tests show that the number of trend trades cannot be reduced below a threshold level, even if the trend calculation period is very long. All trends seem to get caught in a sideways price pattern that generates a series of small losses.

A breakout system, the other very popular trending technique, exchanges a much larger risk for much greater reliability than moving average or exponential smoothing. When a new long position is entered at a new high price for a specific time period, the risk is defined as the price needed to make a new low for the same period. Therefore, the trade risk at the time of entry is the difference between the highest and lowest closing prices, or the highest high and lowest low for the period, n,

$$Trade\ risk = max(High,\ n) - min(Low,\ n)$$

Risk-Adjusted Returns

In order to choose one trading method over another, it is necessary to test them over the same time period and risk-adjust the returns. For example, if system A had an annualized return of 15% and an annualized risk of 10%, while system B had returns of 25% and a risk of 16%, we would multiply the returns of system A by 16/10 or 1.6 to get an adjusted return of 24%. Then system B would be slightly better. If you are not interested in the size of the returns then finding the information ratio, dividing the returns by the corresponding risk, 1.5 for system A, and 1.56 for system B, would give you the normalized relationships.

Time to Recovery

A trading program that recovers quickly from losses is more desirable than a program that lingers at its lowest level of loss. The fact that a program recovers quickly does not change its level of absolute risk (maximum drawdown), but it will change the annualized volatility, measured using a standard deviation, or any statistic that averages the daily negative excursions.

Stability of Returns

The most common calculation of returns is the average of the individual annualized returns; therefore, if a trading program returned 24%, 18%, 20%, 15%, and 19% in 5

consecutive years, most investors would see this as an average return of 19.2%. Another equally valid method is to find the slope of a straight line through the cumulative returns, 24%, 42%, 62%, 77%, and 96%. This gives the annualized returns as 17.9% because it reflects the slight tendency for returns to decline.

A systematic way of determining if your system worked better in the past but has been struggling in more recent markets is to find the slope through the annualized returns (monthly returns if much less data). If the slope is rising or sideways, then the system is performing well. If the slope is noticeably declining, then performance is decaying; it is not adjusting to current or changing markets.

How Much Bigger Could the Risk Really Be?

Many of the risk measurements described in this section attempt to express the probability of a future loss. Within reason, many of them succeed, with VaR being the most likely. An investor must distinguish between the calculated probability and the effects of a real loss. If there was only a 1% chance of a 50% loss in a single month, then by the 50th month there would have been a 50% chance of a 50% loss. There is also a large error factor in most of these calculations because the amount of data used to find the statistics was too small. If 100 cases were used, there is a 10% chance of error in the results. The value of the measurement is also dependent upon the period in the market that was used to create these values; the risk in crude oil was exceptionally high during the Gulf War and then again in 2007. If those periods are not included, based on the assumption that they were outliers, the results will not reflect the future risk of trading crude oil.

It's Not the Markets, It's the Money

At the risk of repeating this too many times, market trends and market shocks are caused by institutions and investors moving money. When all markets move together it is not the fundamentals, it is simply panic. Under stress, investors withdraw money from every type of investment, causing them all to reverse and the correlations to go to 1.

Money seeks a safer place during times of extreme uncertainty, and this is seen as stock prices drop on widespread selling and bond prices rise on widespread buying. In 2011 and 2012, short-term rates went negative—investors were willing to pay the government to hold their money just to gain safety. No amount of correlation analysis or diversification will protect an investment from a price shock in which all markets move the same way based on a flight to safety. No amount of insightful analysis would have allowed for interest rates to go negative.

LEVERAGE

Leverage is controlling an asset using less money than the face value of that asset (also called the *exposure*). If you bought a gold futures contract on 10% margin, then you are depositing only $16,000 to control the value of an asset worth $160,000, 100 ounces of

gold at $1600/ounce. Then every price changing in gold, up or down, is magnified 10 times. If gold drops $10/ounce, then you make or lose $10 × 100 ounces = $1000. In the current volatile environment, gold can easily move $25 each day, or $2500, over 15% of your deposit. That is great when it's a profit, and horrible when it's a loss.

The consequences of leverage are readily seen by measuring the volatility of returns, and the consequences of high and low leverage will be discussed; however, leverage also plays a crucial role in both the trading strategy and the portfolio structure. Futures and options markets offer exceptionally high leverage opportunities, and some traders and analysts act as though they are obligated to take advantage of the maximum allowable. Without leverage, many futures prices show less risk than stocks. In stocks, leverage is only achieved by borrowing up to 50% of the face value of the stock, and that works best when interest rates are low. For the most part, the discussion in this chapter applies to futures. Leverage also allows you the flexibility to balance the risk of a portfolio, a subject that is discussed in more detail in the next chapter.

Rules for Trading on Margin

A few rules should be understood at the beginning. Trading in futures is highly risky because of the leverage. Each exchange sets the *minimum margin* based on price volatility, but the investment house, or broker, often sets that number higher for their smaller customers. In addition, they may require a minimum deposit. They need to protect themselves in the event that a price shock wipes out the entire value of their customer's account and they are left making up the loss.

Commercials, those businesses dealing in commodities, can often put up only 5% rather than 10%. In some cases they substitute a letter of credit rather than any cash. In addition, they *sweep the account* each evening, removing any excess profits and replenishing any losses. There is no audit of how the positions taken by commercials offset business risk, so it is very common for these businesses to speculate under the guise of hedging. When the CFTC publishes its Commitment of Traders report, separating the open interest into speculators and commercials, those who believe that the hedge positions are offsetting physical exposure might be surprised.

Margin Rules and Volatility

The rules of brokerage houses require that an investor replenish the amount in his or her account back to the full amount if the loss on any futures position exceeds 25% of the initial margin. If you were required to put up $16,000 for one gold contract, and gold fell $40, you would get a margin call that required another $4,000 deposited by the next day. If you capitalized the account at more than the initial margin, which is the usual process, then that $4,000 can be taken from the excess in the account. Because $40 is a move of only 2.5%, it is very impractical to have only $25,000 in an account that trades 1 contract of gold. If you started with $50,000, then your leverage is not really 10:1, based on margin, but only about 3:1, based on the entire investment. The minimum funds required by your broker is negotiable, based on your financial solvency and the size of the positions you trade, but can never be less than the minimum exchange margin.

Volatility will vary across markets, and the margin does not always reflect the difference in risk. Exchanges do not change the margin requirements often, and volatility must increase or decrease substantially, and remain at that level, before the exchange governors meet for their review.

Reserves and Targeted Risk Levels

How much do you really need in your trading account to be "safe" and still leverage your returns? During the past 10 years, the hedge fund managers and CTAs have gotten much more conscientious about identifying risk and figuring out the leverage needed to target a particular level of risk. Using the daily profits and losses from a simulated trading history, the process goes as follows:

1. Record the daily profits and losses.
2. Calculate the standard deviation of those profits and losses.
3. Decide your *target volatility*. That is the amount of risk you are willing to take, expressed as annualized volatility. Typically, this is about 15%.
4. Divide the standard deviation of returns (2) by the target volatility to get the investment size needed to achieve that volatility. If the investment is too big by say 25%, then your positions sizes are too big for the investment. Reduce the size of all your positions by 25%.

As an example, Table 23.1 shows a segment of daily returns. Column 2 (*P/L*) shows the daily profits and losses trading 1 contract of EURUSD futures using a trend system. At the bottom on column 2 is the annualized standard deviation (the standard deviation $\times$ $\sqrt{252}$). Using a target volatility of 15%, we divide the annualized standard deviation by 0.15 to get the initial investment needed for that volatility, $266592, which starts column 3. Each P/L change is then added to the initial equity to get the running equity totals. The returns, in column 4, are the P/Ls in column 2 divided by the previous equity value, so that a P/L on 7/26/2011 uses the equity from 7/25/2011. Finally, the NAVs are calculated using $NAV_t = NAV_{t-1} \times (1 + r_t)$. When the annualized standard deviation of the returns is calculated at the bottom of column 4, it comes very close to the target of 15%.

It may seem surprising that an average daily P/L of $1968 requires an investment as large as $266,592, 135 times the average, but one standard deviation of this sample is nearly $40,000, and we can expect to lose at least 2 or 2.5 standard deviations over time, or $100,000. That would be 37% of the initial investment. Then 15% volatility is an aggressive target.

Reserves are the excess funds in the account not used for margin or committed for other purposes. Reserves are needed to cover losses and continue to trade at the targeted level of volatility. Reserves will vary as the position sizes are adjusted in order to keep volatility stable. Most investors, as well as professional managed accounts, set a maximum risk, at which point the positions are liquidated. The maximum level of risk is usually

TABLE 23.1 Calculating the Investment Size Based on Target Volatility

Date	P/L	Equity	Returns	NAV
7/19/2011	0	266592		100.00
7/20/2011	−486	266107	−0.0018	99.82
7/21/2011	−1155	264952	−0.0043	99.38
7/22/2011	1692	266643	0.0064	100.02
7/25/2011	1248	267892	0.0047	100.49
7/26/2011	537	268428	0.0020	100.69
7/27/2011	−198	268230	−0.0007	100.61
7/28/2011	−198	268032	−0.0007	100.54
7/29/2011	1750	269782	0.0065	101.20
8/1/2011	−1108	268673	−0.0041	100.78
8/2/2011	2648	271322	0.0099	101.77
8/3/2011	2544	273865	0.0094	102.73
8/4/2011	−863	273002	−0.0032	102.40
8/5/2011	−852	272150	−0.0031	102.08
8/8/2011	7164	279314	0.0263	104.77
8/9/2011	3477	282791	0.0124	106.08
8/10/2011	4819	287609	0.0170	107.88
8/11/2011	−3827	283782	−0.0133	106.45
8/12/2011	−1038	282744	−0.0037	106.06
8/15/2011	1797	284541	0.0064	106.73
Ann StDev	39989		14.5%	
Target Vol	15%			
Investment	266592			

50%, but it can be as low as 10%. In the previous paragraph, 2.5 standard deviations of the volatility gave a 2.5% chance of a loss greater than 37%, which would be reasonable with a maximum drawdown of 50%. However, a 10% maximum would be equivalent to a target volatility of $^1/_5$ of 15%, or 3%, which would also reduce the expected returns by 80%. In the real world, the lowest practical volatility target is 6%, more often 8%.

LEVERAGE BASED ON EXPOSURE

The traditional way to measure leverage uses *exposure*, the actual value of goods purchased. For example, the margin for 1 gold futures contract was raised to $9,450 on August 24, 2011, when gold was trading at about $1,800. That is 5.25% of the contract value, or leverage of 19:1. As mentioned before, a brokerage firm would never allow a customer to open a futures account with $10,000 to buy or sell 1 gold contract. A move of $25, only 1.4%, which happens often, would create a gain or loss of $2,500. If a loss, then the account value would fall to $7,500, generating a margin call and requiring replenishing the account by the next day. Each brokerage firm decides on the minimum account value

based on what will be traded, as well as some assessment of the experience and credit history of the trader. For 1 contract of gold, it is likely that the broker will require $50,000, reducing the leverage to 3.6:1, a more likely scenario.

In Europe, some registered funds have leverage limits set by the government. Typically, this can be 2.5 times the current account value. Compared to U.S. regulation, this is very severe. For the gold contract valued at $180,000, the account must set aside $72,000. The biggest problem is with interest rate futures, which are also assessed at face value, although the risk is actually a function of the yield, which goes in the opposite direction. Then when Eurodollar rates are trading at 99.50, yielding only ½% interest, the regulatory authorities require $0.9950 \times \$1,000,000$ (face value), or $995,000 to trade one contract. That essentially eliminates trading Eurodollars in a small investment account because, as we will see in the next section on the calculation of position size, the current low volatility of Eurodollars requires a large position compared to other markets in a diversified portfolio. This risk calculation also causes most funds to choose a different method for calculating risk, such as VaR, which is permitted by European regulators.

INDIVIDUAL TRADE RISK

Risk control begins at the moment a position is entered. There are two distinct issues at that time, the strategic element of the trade, for example, the risk and reward of a single share or contract, and the total exposure of that trade relative to the investment amount. We will look at the second one first.

Position Sizing

The traditional way of deciding how many shares to trade in a stock is to divide the total investment by the number of stocks in the portfolio, then divide that portion of the investment by the share price. Then each stock has the same invested amount, but not the same risk. For example, if we were trading four stocks, Apple, Amazon, Microsoft, and Bank of America where each stock had a nominal risk amount of $10,000 on October 7, 2011, the position sizes (line 2 in Table 23.2) could be found as $10,000 divided by the share prices (line 1). Then Bank of America has a very large position, 1695 shares, compared to the other stocks. But it turns out that BAC, trading at a low price, is relatively more volatile than the other stocks and this position size focuses all the risk on BAC.

There are three basic ways to calculate the position size of any trade in stocks, given $10,000 to invest:

1. $10,000 divided by the share price
2. $10,000 divided by the dollar value of $100 \times$ annualized volatility
3. $10,000 divided by the average true range $\times \sqrt{252}$

TABLE 23.2   Comparison of Position Sizes Based on Three Methods

7-Oct-11	AAPL	AMZN	BAC	MSFT
Close	369.80	224.74	5.90	26.25
10000	27	44	1695	381
Ann Vol	424	246	134	345
AvgTR	71	66	1400	2033
	AAPL	**AMZN**	**BAC**	**MSFT**
10000	62.7	38.1	1.0	4.4
Ann Vol	1.0	1.6	62.7	14.1
AvgTR	1.1	1.0	21.1	30.6

Table 23.2 shows the considerable variation in the position size using these three methods. The upper part of the table gives the number of shares, and the lower part normalizes the shares by dividing each value on that row by the lowest value on the row. Then method 1 ("10000") shows that the low price of Bank of America results in a position size 62 times the size of the Apple position. In method 2, the annualized volatility, the results of Bank of America are reversed because the percentage returns of a volatile stock trading at a low price are unusually high. Instead of having the largest position, as in method 1, it has the smallest. Method 3 uses the highs and lows to find the average daily risk, then annualizes the value. This time Amazon has the highest risk and Microsoft the lowest. The three methods produce astonishingly different results.

The problem with calculating size for equities using a method other than (1) is that dividing the funds equally among the stocks does not work. Using method 1, the maximum number of shares for AAPL would be 27, so that there would not be enough money to buy 424 shares required by method 2, or 71 shares for method 3. It will be necessary to pool all of the money and redistribute it in proportion to the number of shares needed. For example, part 1 of Table 23.3 shows the dollar value needed to trade each of the stocks in each method given by Table 23.2. On the far right is the multiplication factor needed to reduce the totals of methods 2 and 3 to the $40,000 available. Part 2 multiplies the values in part 1 by the factor on the respective line, so that the investment required for each method comes to the same $40,000. Part 3 then divides the respective investment amounts from part 2 by the share price to get the actual shares traded.

Methods 2 and 3 in part 3 are fully risk-adjusted, but show a significant difference in the allocations. In method 2, there would be very low exposure to BAC, more to AAPL, and less to MSFT, while method 3 is very different. Which is better? The true range method is considered a better measure of risk; however, it would be best to test each method over a wide range of stocks and futures and use the average information ratio to draw your own conclusions.

As an example, we will compare the traditional method 1 with method 3 using BAC because it had the greatest variability, and apply a 200-day moving average strategy over the past 20 years. Before discussing the results, it is important to understand

TABLE 23.3   Adjusting the Shares Traded for Each Method of Position Sizing

1	Investment amount needed to trade each method						
Method	**AAPL**	**AMZN**	**BAC**	**MSFT**	**Total**	**Factor**
10000	10000	10000	10000	10000	40000	1.000
Ann Vol	156784	55331	793	9048	221956	0.180
AvgTR	26241	14922	8262	53357	102782	0.389

2	Total investment redistributed among stocks according to method					
Method	**AAPL**	**AMZN**	**BAC**	**MSFT**	**Total**
10000	10000	10000	10000	10000	40000
Ann Vol	28255	9972	143	1631	40000
AvgTR	10212	5807	3215	20765	40000

3	Shares actually traded after redistribution of investment				
Method	**AAPL**	**AMZN**	**BAC**	**MSFT**
10000	27	44	1695	381
Ann Vol	76	44	24	62
AvgTR	28	26	545	791

that a single test cannot be conclusive because the best method should work regardless of the calculation period or the strategy. First, the total P/L, shown in Figure 23.5, indicates that BAC is not a successful candidate for trend trading, but the average true range method shows far less volatility, 27% lower, than trading the same shares for each trade (this test used 2201, which was the average shares traded using the true range method). Figure 23.6 shows a history of the number of shares traded, with
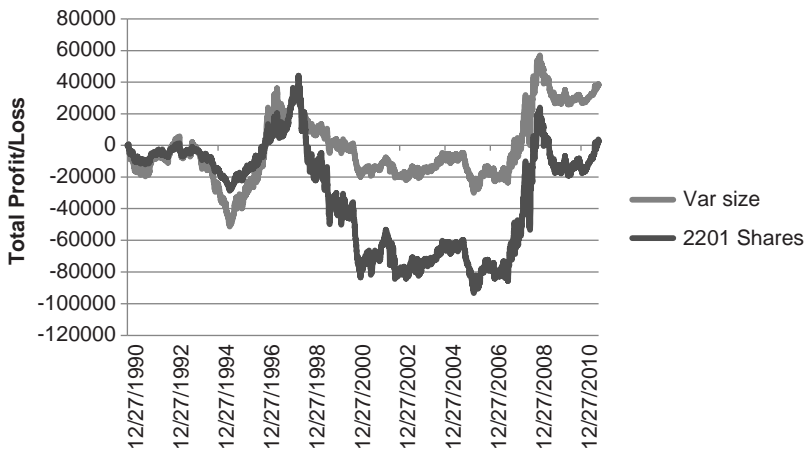
FIGURE 23.5   Total PL for BAC using methods 1 and 3, for 20 years.
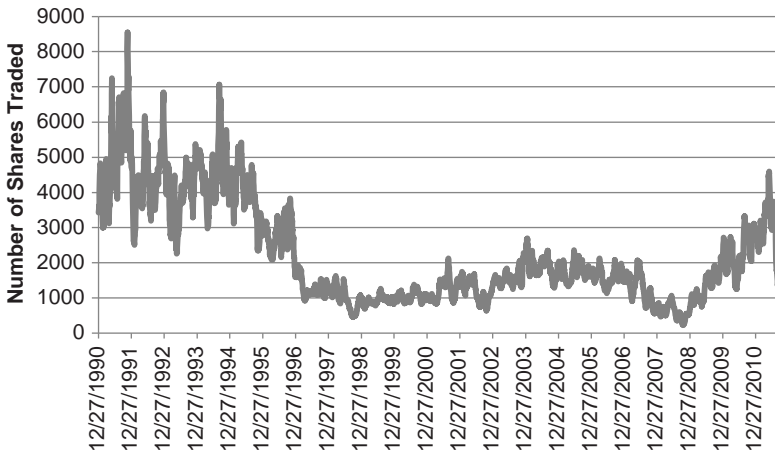
FIGURE 23.6 Number of shares traded in BAC using the true range method.

exposure dropping sharply from 1994 to 1996 while prices were moving steadily higher (see Figure 23.7). Position sizes dropped from about 6000 shares to 1000 shares. The exceptional volatility in 2008 did not cause as much of a reduction in position size, falling from about 1500 shares to 500 shares. Be reminded that varying the number of shares based on volatility constantly changes the amount needed as an investment, so that a rigorous test would need to define the investment size as the largest amount needed over time.
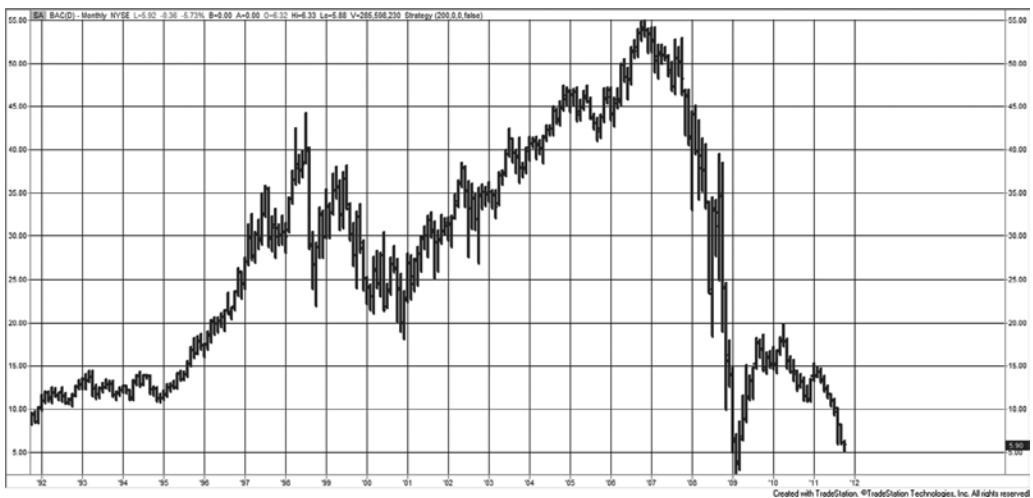
FIGURE 23.7 Bank of America, monthly from 1992 through September 2011.

Even with all the qualifications and calculations, adjusting the position size to target the same risk on each trade gives each trade an equal chance to participate in the results. Unless you can select which trades are most likely to be better than another, equal risk is the most conservative approach. Reducing the position size when the market gets more volatility also prevents the volatility of returns from varying wildly. The price swings of the past 3 years would easily have resulted in portfolio volatility ranging from very low at 6% to as high as 25% if no attention was paid to position size.

Initial Position Risk

The entry point of a trade is the time of greatest uncertainty. Systematic strategies identify the critical point where the trend should change, but when the calculation uses a large number of days, the lag may cause a delayed reaction. In the case of a moving average system, the trend will converge on current prices during a sideways pattern, at which point the trend direction may flop back and forth until prices move out of the range. The period of greatest trend uncertainty is the point at which it changes direction.

For mean reversion traders, the greatest profits are in anticipating the change in direction of prices. They must buy when prices are falling, clearly an uncertain time. They count on the likelihood that certain moves cannot be sustained for more than a few days or that large moves are overblown. There is considerable uncertainty at the point of entry.

Each trading style has its own intrinsic risk associated with initial entries. The trend philosophy is known for its tenacity, entering a long position whenever prices turn up, taking a small loss if prices then turn down, and reentering the next time they turn up. The breakout technique will limit losses to the size of the current trading range and any attempt to reduce this loss, which can be large in a volatile market, conflicts with the essence of the method. Mean reversion systems must take large losses in exchange for a high percentage of smaller, profitable trades. If losses are limited using stops, then the percentage of profitable trades is reduced, and the profit/loss balance is upset.

Stop-Losses

Stops provide the ability to control risk on individual trades. A *sell stop* is placed below the current price and becomes a *Market Order* to sell once that price is touched. Stops do not guarantee that you are filled at the stop price because opening gaps, fast markets (after an economic report), and low volume can result in large slippage. If the market gaps lower on the open, below your stop order to sell, the order gets filled at the market price. Stops placed close to the price currently being traded are hit frequently. Because of the nature of price movement, the farther away you place a stop, the less often it is hit. The pattern is very similar to a normal distribution. Using random numbers, the number of times reached multiplied by the size of the stop will always be constant. Then a stop must be better than a random event to be a good trading rule.

The beginning trader may decide that no trade should lose more than $500, and place a stop at that point. Some more experienced traders believe that you should never lose

more than 3% of your total investment on a single trade. On some level, it makes sense to scale your losses to the amount of money at risk, but stops are more sensible if they relate to the nature of the system, the timing of the trade, the market volatility, or a chart pattern. Some worthwhile possibilities that change with market conditions are:[12]

1. Advance the stop by a percentage of the price change, as in Wilder's parabolic system.
2. Use a swing high or low point, based on a percentage minimum swing.
3. Use the highest high or lowest low of the recent n periods.
4. Apply a method such as Kaufman's *Adaptive Moving Average* (KAMA) as a stop.
5. Adjust the stop by the volatility, such as 3 times the current 10-day average true range.

Initial Stop

Because the greatest period of uncertainty is at the point of the trade entry, a stop can be placed at that time and never moved. But trend systems already cut losses quickly, and mean reversion systems need a high percentage of profitable trades, the result of holding trades that initially go the wrong way, so it is not clear that any stop-loss would improve those profiles. Under highly volatile situations, they may be a benefit, but that would need to be tested for specific strategies, and the frequency of occurrence would be small.

Trailing Stops

A *trailing stop* is one that captures an increasing part of the profits as prices move in a favorable direction. Some trailing stops advance, but never retreat, as in Wilder's *Parabolic*; others, based on volatility and price changes, may retreat. Trailing stops are usually constructed in one of the following ways:

1. *Fixed percentage.* A long position was entered in IBM at $75 using a moving average. A trailing stop of 3% is calculated based on the highest prior close. The trade is liquidated if either the trend changes or prices fall by 3% from the high point of the trade, whichever comes first. Alternatively, a stop may be a percentage of the trend value and placed at the trendline minus 3% of the trendline, forming a smooth band.
2. *Volatility.* Because price volatility changes, a trailing stop may be calculated using any of the volatility measures discussed earlier, preferably annualized volatility or average true range, and placed below the highest high of the current price move. This method is more adaptable than the fixed percentage; however, when volatility increases, today's stop may move farther away than on the previous day. ISAM has suggested that a generic stop is $12 \times average\ true\ range$ calculated over the past 252 days; however, this conflicts with the idea that the multiplier should vary based on the calculation period of the trend.

---

[12]Tushar Chande and Stanley Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994).

3. *Percentage of profits.* Placing a stop at a point that captures 50% of the profits is a sensible technique except at the beginning of the trade. When there are only small profits, the stop would be too close; therefore, this method can only be used after some profits have already been accumulated in the trade, a *trigger* value. Prior to that, an initial stop would be based on price volatility. Once the profit stop is triggered, the closer of the two stops would be used.

Standard Deviation Stop

The standard deviation is the basis for traditional risk assessment, and it can be used to determine stop-loss levels.[13] In her *Dev-stop*, Cynthia Kase uses the following steps to create stop-loss levels for both long and short positions:

1. Calculate the *true range* (TR) of the past two trading days using the highest high and lowest low of the 2-day period.
2. Calculate the rolling *average true range* (ATR) of TR (in Step 1), using 30 periods for intraday charts and 20 periods for daily charts. Note that each period combines 2 days.
3. Calculate the standard deviations (STDEV) of the true ranges in Step 1 using the same 20- and 30-day period as in Step 2.
4. The stop-loss values are *DDEV = ATR + (f × STDEV)*, where *f* = 1, 2.06 to 2.25, and 3.20 to 3.50, where the larger values of the pairs correct for skew and allow for greater risk.
5. The *dev-stop* for long positions is *trade high – DDEV;* the dev-stop for short positions is *trade low + DDEV. Trade high* and *trade low* are the prices corresponding to the greatest profit in the long and short trades.

This method adjusts for volatility using a standard statistical measurement and is applied to the extreme profit of a trade, in the manner of a trailing stop. In Figure 23.8, an 80-day moving average, the smooth, heavier line, is shown with 30-year bonds. The two other lines above and below prices are the standard deviation stops based on the 80-day price differences and a multiplication factor of 8. That value is then subtracted from the highest price of a long position or added to the lowest price of a short sale. The chart shows that the trend change from long to short, which occurred about December 1, 2010, would have been stopped out at least 15 days earlier. In the following short position, the daily high penetrates the stop, but the stop is only activated if prices close above that level. In the most recent long position, the stop is closer to prices than the moving average but manages to stay clear of being hit because it adapts to the price volatility. Two programs that use these stops and display them in an indicator are *TSM Moving Average Stop-Loss* and *TSM Stdev Stop-Loss*, both on the Companion Website.

---

[13] Cynthia Kase, "Redefining Volatility and Position Risk," *Technical Analysis of Stocks & Commodities* (October 1993).
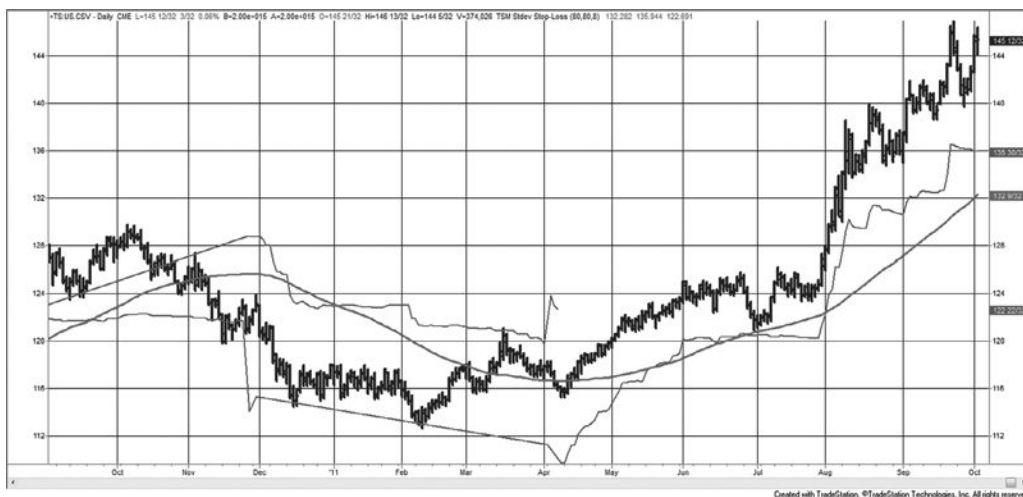
FIGURE 23.8 30-year bonds with an 80-day moving average and an 80-period standard deviation with a factor of 8 measured from the high or low of the trade.

Profit Targets

Profit targets are points at which the system captures a profit expected to be larger than the return at the end of the trade, or a short-term windfall, such as a favorable price shock. "Profit taking" may actually net a loss on the trade if a sudden partial recovery from a larger loss looks attractive. Profit targets are essential in short-term trading, where the price noise makes it likely that any favorable move will reverse soon. They are more difficult to incorporate into a longer-term trend-following system because they risk missing the bigger profit (from the fat tail) in exchange for smaller but more frequent profits. As a trend-follower, once you take profits but the trend direction is still intact, you need to find a way to reenter that trade to avoid missing the rare, but very large price move.

Profit-taking is best when it is based on price volatility (V), and is most often calculated as

1. For longs, a multiple of the average true range or annualized volatility added to the system entry point (not the actual fill), $E + f \times V$.
2. A favorable move of $f \times V$ on any single day, or combination of successive days, where f is a multiplication factor.

While any profit target assumes that prices will not continue in the same direction, method 2 may not net a profit on the trade. A large, favorable move may occur after an initial trade loss, and the net return may still be a loss. However, if the extreme move cannot be sustained, then taking a small loss is preferable over a larger loss.

The factors used for the net size of the profit-taking level are usually arrived at by data mining. Not to be outdone, we will look at the various levels of profit-taking applied to a short-term and longer-term moving average system for 60-minute euro currency futures.

In Table 23.4 the 60-minute euro futures are shown with an 80-bar and a 5-bar moving average system where profit-taking has been applied using a factor (shown on the left) multiplied by the average true range over a 20-bar period. Once the profit target has been hit, the system waits for the next trend trade. It cannot reenter the same trend. The top line shows the results with no profit-taking. The longer 80-bar average has the smallest profits but the highest percentage of profitable trades when the profit target is smallest. As the target gets larger, profits get larger. The no-profit case at the top shows that returns were best if there was no profit-taking.

The faster trend on the right shows that profit-taking turned a losing strategy into a profitable one for factors from 5.50 to 9.0. They were not large profits, but because of the higher level of noise in 5-minute bars, profit-taking can take advantage of short-lived price moves. Because longer calculation periods emphasize the trending nature of prices, taking profits prevents the system from realizing all of the trend.

A program to test variable-size profit targets, *TSM Moving Avg Profit Targets*, can be found on the Companion Website.

TABLE 23.4 Profit-Taking Tests Using the Euro Currency Futures, 2002–2010

PT Factor	80-Bar Moving Average				5-Bar Moving Average			
	Net Profit	% Profitable Trades	Avg Trade	Profit Factor	Net Profit	% Profitable Trades	Avg Trade	Profit Factor
None	121655	45	1580	2.738	−3278	32	−7	0.991
10.00	94405	45	1226	2.348	−3278	32	−7	0.991
9.50	85693	45	1113	2.224	−3278	32	−7	0.991
9.00	90880	45	1180	2.298	4723	32	11	1.013
8.50	86030	45	1117	2.229	4160	32	9	1.012
8.00	81630	45	1060	2.166	6498	32	14	1.018
7.50	75543	45	981	2.079	19098	32	43	1.053
7.00	68168	45	885	1.974	11485	32	26	1.032
6.50	71483	46	916	2.021	6635	32	15	1.019
6.00	63583	46	815	1.908	4448	32	10	1.012
5.50	53945	46	692	1.771	523	32	1	1.001
4.50	35533	46	456	1.508	−11915	32	−27	0.967
4.25	30608	46	392	1.437	2785	32	6	1.008
4.00	26345	46	338	1.376	−8728	32	−19	0.976
3.50	17583	46	225	1.251	−14153	32	−32	0.960
3.00	25383	49	325	1.385	−27478	32	−61	0.923
2.50	21033	51	270	1.324	−29538	32	−66	0.917
2.00	22208	54	285	1.363	−31050	34	−69	0.911
1.50	13658	56	175	1.233	−24525	38	−55	0.927
1.00	18220	62	234	1.448	−40113	43	−89	0.873
0.50	11308	69	145	1.306	−37775	54	−84	0.862

Multiple Target Levels, Scaling Out of a Trade

Floor traders instinctively know how to reduce risk and give themselves the best opportunity for profit.[14] One technique that stands out is the way they take profits. Because they are dealing with trades that may last from minutes to hours, they are sensitive to market noise and realize that the size of the price move is very unpredictable. Instead of targeting a single price level as an exit point, they will often exit in three parts. For example, if you bought the *e*mini S&P at 1125 and with an average daily volatility of 15 points, you might expect to target an exit of 1133. But what if the price reached 1132? Floor traders cannot count on being rigid about a target, so they are more likely to set three targets 1130, 1133, and 1135 to cover a wider range. If the first target is hit, they have reduced their risk considerably. If the second target is hit, they will never turn a profit into a loss. Then multiple targets, especially where the average is the same as your original target level, reduce your risk as the market goes yours way, and still give you a reasonable chance to capture the same total profit.

KAUFMAN ON STOPS AND PROFIT-TAKING

Both profit-taking and stop-losses are a duel with price noise. If volatility causes a sharp drop followed by a recovery, then hitting a stop captures the worst exit. On the other hand, if volatility gives you a windfall profit, then exiting that position will be a good move. Because there is so much market noise, stop-loss orders are usually based on the closing price, or, if they are triggered by intraday price moves, the actual exit is still on the close to take advantage of a pullback. The opposite would be true of profit-taking. You would want to exit at the time of the intraday spike rather than waiting for the end of day.

There is a lot of debate about whether a stop-loss improves results or simply seems to be a good idea but does not actually result in better performance. Every trader can tell a story of how a stop order was executed at the worst price of the day, or how it saved him from a devastating loss. In fast markets, you are lucky if your stop is executed near your price. Often it gets filled at the extreme of the move, where all the buyers and sellers cluster. But that does not mean that stops do not work. Some work, and some do not work. The ones that are most likely to work must adapt to volatility, such as the standard deviation stop, rather than a fixed dollar amount or a percentage of price. And trailing stops are more practical than initial stops.

As an example, a test was run on U.S. 30-year bonds from 2000 to 2011, using a moving average trend strategy with calculation periods from 10 to 100 in increments of 10 days, and a standard deviation stop based on a fixed calculation period of 40 days but varying the factor from 2 to 10 in increments of 2. This produced 50 tests. Of course, we

---

[14]Courtesy of Barbara Diamond, the first woman trader on the floor of the Chicago Mercantile Exchange.

TABLE 23.5   Stop-Loss Test of 30-Year Bonds, 2000–2011, Applied to a Moving Average Strategy and Displayed as a Heat Map

MA Period	Stop-Loss Factor					Average	No Stop
	10	8	6	4	2		
100	–11594	–6969	8625	11594	–6313	–931	–27875
90	10063	12250	35781	47594	35438	28225	688
80	11625	13063	33594	44688	21688	24931	4813
70	26063	32219	42406	41406	34063	35231	8531
60	1688	7844	13813	14250	16906	10900	6281
50	–10281	–1156	12750	22313	–344	4656	–8000
40	5906	13375	21375	30906	20156	18344	–3781
30	–19000	–13531	–3344	6781	7125	–4394	6938
20	33938	36156	34125	52500	44688	40281	36375
10	–1594	–813	1094	–188	17438	3188	–813
Average	4681	9244	20022	27184	19084	16043	2316

could have varied the calculation period for the stop and made the increments smaller, but then you would expand the number of tests considerably. This is only to give us the big picture.

Table 23.5 is a *heat map* arranged to see the continuous patterns. Although the tones may not be clear here, you can create the same chart using the *conditional formatting* option in Excel. The left scale is the moving average period, and the top is the stop-loss factor. The upper left has the slowing moving average and the largest stop, and the bottom right is the most sensitive. The right-most column gives the profits of the same moving average strategy with no stop. There are also averages of the columns and rows. In comparing the two columns on the right, every stop-loss average except one was better than the corresponding strategy with no stop-loss. On the bottom, every stop-loss factor was better than the average of all trends with no stop loss. By any standard, we would consider this a robust improvement. Keep in mind that this is a test of only one market and over one selected time period. Interest rate markets are well known to have trends, so that results might be very different applied to equity index markets that are typically noisy.

Risk Protection or False Hope?

Stops are normally *resting orders*, that is, they are held by the floor broker or in the electronic queue, to be executed as a *market order* when the price is touched. When a price shock occurs and the market jumps through your stop order, it is very likely that you will get filled at the worst price of the fast move. So stops work as you planned most of the time, except when you need them most. It is very similar to other financial phenomena, such as correlation, where a portfolio normally benefits from diversification except during times of extreme stress.

Market Noise, Trends, and the Frequency of Stops

Because of market noise, stops that are close to the market price will get hit more often. As discussed briefly before, if we were dealing with random price data, the number of times that a stop is hit, times the distance of the stop from the market, will always be constant. Then a stop that is 100 points away may get hit 10 times for a total loss of 1000 points, and a stop that is 200 points away will get hit 5 times, also for a total of 1000 points. However, price data is not random, and the distribution implies that larger moves are more likely to change the direction of the trend. As you filter out more of the noise, those remaining events have a greater likelihood of being significant.

Proximity Risk

When many stop orders cluster around the same point, there is *proximity risk*. One way this can be seen is in a gap as price breaks through a key resistance level. Given the large number of trend-following systems, it would not be surprising to find that many of them are generating the same orders on the close, especially following a very volatile price move.

Large macro trend systems, those with substantial assets under management, may take up to 3 days to enter a trade. The likelihood of proximity risk is greatly increased, and the cost of execution could be significant.

Managing Risk without Stops

A stop-loss is not the only way to reduce risk. An alternative is to reduce the size of your positions as volatility increases. This can be done at different places within the trading process. At the beginning of this chapter, we showed how to calculate the position size based on price volatility. The greater the volatility, the smaller the position. You can also rebalance those positions' sizes as the volatility changes, although you must consider the added switching costs. That could be done using a testing platform as long as you correctly estimate and include the commissions and slippage each time you change position size. Volatility stabilization will be discussed in the next chapter.

Risk Control Overlay

Some traders include stop-loss orders based on factors other than the price or volatility of the specific market traded. They generally take one of the following three approaches:

1. *A percentage of initial margin*, for example, 50% to 70% of initial margin. This is loosely related to long-term volatility but lags considerably.[15] An estimate of long-term volatility may be a more satisfactory alternative.

2. *A percentage of the portfolio value or total account value*, for example, 1.0% to 2.5%. This concept of equalized risk (and perhaps reward) across all markets is very

---

[15]Tushar Chande and Stanley Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994).

popular; however, it is not sensitive to individual markets, and as with many stops, it imposes artificial overrides. If the volatility of one market increases or decreases faster than others, then this risk level would need to be adjusted to individual markets; otherwise, it could be reached on every trade or not reached at all.

3. *The maximum adverse excursion* determined by historic evaluation.[16] A stop is placed just beyond the maximum adverse excursion for each trade, or 2.5% of the price, whichever is smaller.

RANKING OF MARKETS FOR SELECTION

Knowing which market to trade at the right time would clearly improve performance. To accomplish this, there are a number of ways to measure the *trendiness* of a market, from standard statistical techniques to more complex rules. Each method will measure some special characteristic of price movement. To take advantage of this, it is necessary to trade using a system that targets this type of price pattern. For example, if you use a correlation coefficient, r^2, to rank the trend over 20 days, then the highest-ranking markets may do best using a linear regression for the trend-following system, with a period no greater than 20 days. If the correlation is low then mean reversion would be favored.

To prove that ranking methods result in better performance, we would need to record their values in a table over various time periods.[17] For example, the five measurements below can be calculated for periods of 5, 10, 20, and 40 days.

1. Correlation coefficient, but only for values greater than 0.25.

2. Sum of the net moves, in dollars, over n, $2 \times n$, and $4 \times n$ days.

3. Slope of an n-day regression converted to a 1-day dollar-value change.

4. Wilder's Average Directional Movement Index (ADX), but only for values greater than 0.20.

5. Average of the absolute value of price changes.

Each measurement will need to have threshold values that determine when they represent significant trends. For the correlation coefficient, anything below 0.25 is most likely to be a sideways market; similarly, Wilder's ADX would need to be above 0.20 to indicate a trend. If the results are erratic, it may be necessary to smooth the last 3 days of each value. In the case of item 5, higher volatility is often associated with greater profits and greater risk for a trend-following system; therefore, it is necessary to take a careful look at how this ranking affects performance.

---

[16]John Sweeney, *Campaign Trading* (New York: John Wiley & Sons, 1997).

[17]Based on an idea suggested in Chande and Kroll, *The New Technical Trader* (New York: John Wiley & Sons, 1994).

Commodity Selection Index

Among Wilder's trading tools is the *Commodity Selection Index* (CSI),[18] a calculation for determining which products are most likely to make the greatest move for each dollar invested. In this case, the movement measured is *directional* and, therefore, should apply directly to trending models. The CSI combines directional movement, volatility, margin requirements, and commission costs into an index that allows for comparison and selection. It is the result of the following calculations.

Directional Movement

The trending quality of the market, as defined by Wilder, begins with *directional movement*, the greater of either:

1. Plus DM (PDM), today's high minus yesterday's high, $H_t - H_{t-1}$.
2. Minus DM (MDM), today's low minus yesterday's low, $L_t - L_{t-1}$.

Note that *plus DM* and *minus DM* are often written as +*DM* and –*DM;* however, to avoid confusion in the following calculations, they will be shown as *PDM* and *MDM*.

The directional movement is either up or down, whichever is larger of PDM and MDM. It is the largest part of today's range that is *outside* yesterday's range. The value that is not used is set to zero. When an inside day occurs, both PDM and MDM are zero, and the directional movement is zero (Figure 23.9).

In the calculation of PDM and MDM, DM is expressed relative to today's *true range* (TR1), $\max(H_t - L_t, H_t - C_{t-1}, C_t - L_{t-1})$. The true range is always positive.

The relationship of the price direction to the true range is called the *Directional Movement* (DM). Today's directional movement is calculated using either the PDM or the
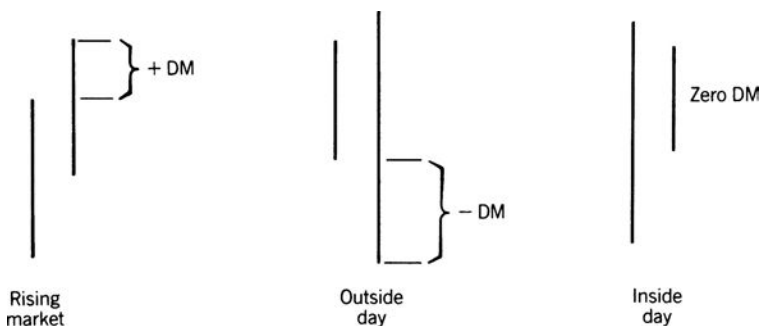
FIGURE 23.9 Defining the DM.

---

[18]J. Welles Wilder, "Selection and Direction" in Perry J. Kaufman, ed., *Technical Analysis in Commodities* (New York: John Wiley & Sons, 1980).

MDM, whichever is greater. The notation "14" refers to the period over which the values are smoothed.

$$PDM14 = +\frac{DM14}{TR14} \quad or \quad MDM14 = -\frac{DM14}{TR14}$$

Once the first DM14 is calculated, using 14 days of price movement, an *average-off* technique is used to find each successive DM14 as follows:

$$PDM14_t = PDM14_{t-1} - \frac{PDM14_{t-1}}{14} + PDM14_t$$

$$MDM14_t = MDM14_{t-1} - \frac{MDM14_{t-1}}{14} + MDM14_t$$

The same procedure is followed for the true range:

$$TR14_t = TR14_{t-1} - \frac{TR14_{t-1}}{14} + TR14_t$$

These results can also be produced using a smoothing constant of 0.071, which is an approximation of the fraction $\frac{1}{14}$, where t is today's value:

$$PDM14_t = 0.071 \times PDM14_{t-1} + PDM14_t$$
$$MDM14_t = 0.071 \times MDM14_{t-1} + MDM14_t$$
$$TR14_t = 0.071 \times TR14_{t-1} + TR14_t$$

At this point, the Directional Movement components can be used as trading indicators; however, Wilder's interest was to use this in a more complete concept, shown in the following sections.

Directional Indicator and True Directional Movement

Once the PDM14, MDM14, and the TR14 are calculated, the *directional indicators*, PDI14 and MDI14 follow,

$$PDI_t = \frac{PDM14_t}{TR14_t}$$

$$MDI_t = \frac{MDM14_t}{TR14_t}$$

and the *True Directional Movement* (DX) is the difference between PDI14 and MDI14. When an upward trend is sustained, the MDI current value is zero; therefore, the PDI14 becomes larger, the MDI14 becomes smaller, and DX becomes greater. This is then *normalized* in order to express the final value between 0 and 100.

$$DX_t = 100 \times \frac{|PDI14_t - MDI14_t|}{PDI14_t + MDI14_t}$$
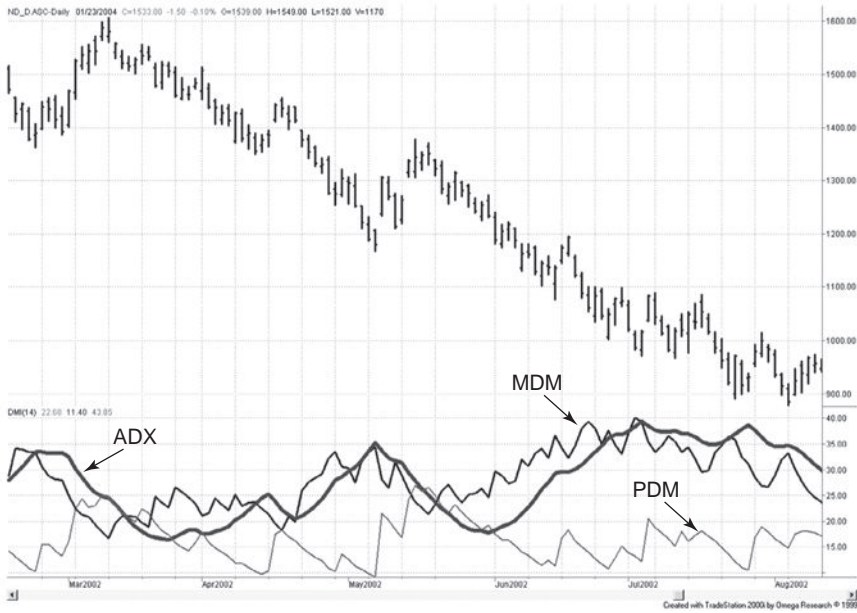
FIGURE 23.10 The 14-day ADX, PDM, and MDM, applied to NASDAQ 100 continuous futures prices, February to August 2002.

where multiplying by 100 converts the percentage to a whole number, and the absolute value prevents DX from becoming negative. At the same time, the absolute value causes DX to lose the information about which direction prices are moving.

Average Directional Movement Index (ADX)

The DX is then smoothed using the equivalent of a 14-day average (a 0.133 smoothing constant), and is called the *Average Directional Movement Index* (ADX).

$$ADX_t = ADX_{t-1} + 0.133 \times (DX_t - ADX_{t-1})$$

The ADX, the PDM, and the MDM are shown for NASDAQ 100 continuous futures in Figure 23.10. The heavier ADX line moves higher as the price trend becomes clear. The individual components, the PDM (shown as the thinnest line) and the MDM (the medium-weight line), indicate the relative strength of the upwards and downwards moves within the trend. As NASDAQ prices decline, the MDM remains above the PDM, which is shown along the bottom of the lower panel.

One last adjustment is made to the extreme variance of the ADX by taking the average of the current value and the value of the ADX 14 days ago. This final value is called the *Average Directional Movement Index Rating* (ADXR),

$$ADXR_t = \frac{ADX_t + ADX_{t-14}}{2}$$
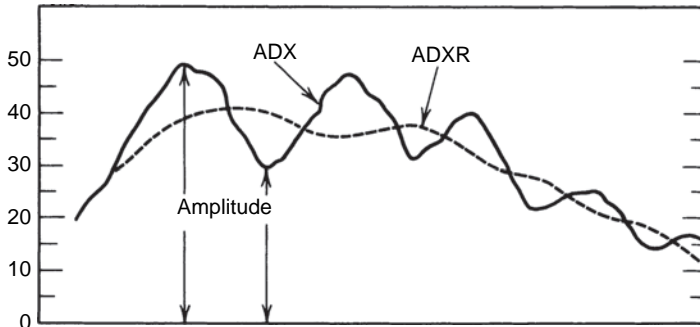
FIGURE 23.11 The ADX and ADXR.

The ADX and ADXR are shown plotted together in Figure 23.11. The ADX is seen to oscillate about the ADXR. Measuring the distance of the ADX from the zero line, a higher amplitude means higher directional movement and a stronger trend, whether up or down. The peaks are always the extremes. The distance between the ADX and ADXR is used to measure the overbought and oversold condition of the trend. The larger the value, the greater the reactions to the trend.

Using the ADX as a Trending Indicator. The ADX has become a popular tool. Ruggiero[19] uses it to determine the trend with the following rules:

1. If ADX crosses above 25, the market is trending.
2. If ADX crosses below 20, the market is consolidating.
3. If ADX crosses below 45 after being higher, the market is consolidating.
4. If ADX rises above 10 on 3 of 4 days after being lower, the market will start to trend.
5. A trend based on rule 4 remains in effect until the 5-day difference in the ADX is less than 0.

Commodity Selection Index

The creation of the Directional Movement Indicator, the different components, and finally the ADX leads to the *Commodity Selection Index* (CSI), which is calculated as:

$$CSI_t = ADXR_t \times ATR14_t \times \left[100 \times \frac{V}{M} \times \frac{1}{150 + C} \right]$$

where $ADXR$ = average directional movement index rating
 $ATR14$ = 14-day average true range
 V = conversion factor; value of a 1¢ move for a futures market (in dollars)
 M = margin (in dollars)
 C = commissions (in dollars)

---

[19] Murray A. Ruggiero, Jr., *Cybernetic Trading Strategies* (New York: John Wiley & Sons, 1997).

The CSI can be applied to equities by setting $V = 1$ and M to the total share value of the investment. Note that for a particular futures market, or for equities, the value inside the bracket does not change. By calculating it once and calling that value K, the CSI can be expressed as:

$$CSI_t = ADXR_t \times ATR14_t \times K$$

A portfolio allocation can be chosen by calculating the CSI daily or weekly for each market. Simply select those products to trade that have the highest CSI or allocate position size in proportion to their CSI value.

Retesting Directional Movement

When these indicators were published in 1978, they were all based on a 14-day calculation period, which is why all of the notations include the number 14.[20] With the ability to test these indicators over different calculation periods, the "14" was dropped from the notation. The first study to use PDI and MDI, the directional indicators, in a trading strategy was by Hochheimer, known for his other studies of moving averages, crossovers, and channels. He defined the rules of a *Directional Movement System* in two ways.[21] The first set of rules is:

1. **a.** *Enter a new long position and close out any short sales* if the PDI crosses above the MDI. Enter a *buy stop* on the next day using today's high price. This order remains as long as it is not executed and PDI remains higher than MDI.

   b. *Enter a new short sale and close out any long positions* if the PDI crosses below the MDI. Enter a *sell stop* on the next day using today's low price. Maintain this order until it is executed and while PDI remains below MDI.

Hochheimer calls the first case "directional movement with delay." The second case is an immediate market entry following the crossing of the directional indicators:

2. **a.** *Enter a new long position and close out any short sales on the open of the next day* if the PDI crosses above the MDI.

   b. *Enter a new short sale and close out any long positions on the open of the next day* if the MDI crosses below the PDI.

In both cases the system is always in the market. Before seeing the actual results, it is possible to generalize the expected performance.

1. Rule 2 must have more trades than Rule 1 because it always takes a position when a crossing occurs, while the first set of rules requires a confirmation.

---

[20] For a general discussion of this topic, see Chapter 21.

[21] Frank L. Hochheimer, *Computerized Trading Techniques 1982* (New York: Merrill Lynch Commodities, 1982).

2. Because there is no commitment to stay in the trade once entered (for example, no price channel), there could be frequent whipsaws using Rule 2.

3. Because Rule 1 uses the high and low of the prior day, its entry prices will always be equal to or worse than those of Rule 2.

4. If the Directional Indicator gives a highly reliable signal, it would be better to enter immediately, as in Rule 2.

In Hochheimer's tests, the following parameter ranges were used:

1. The PDM varied from 7 to 20 days.

2. The MDM varied from 5 above to 5 below the PDM value.

3. Two *true ranges* (TR) were calculated, the first using the parameter value of the PDM, the other using the parameter value of the MDM.

The test data covered a wide range of those futures markets available from 1970 to 1981.

Rather than show the detail of those tests, which would be considered out of date with current markets and conditions, we ran a simpler test on directional movement using the first set of rules, varying only the basic calculation period. Tests include five most popular markets, one from each sector, and covered from 2000 through 2011. Results are shown in Table 23.6. The calculation periods were from 4 days to 80 days in increments of 4 days. While each market has areas of good profit, there was little consistency. The best period for bonds was 4, the DAX 54, EURUSD 28, crude 20, and copper 60. This alone does not satisfy our conditions for a robust system.

Trading Rules Combining PDM, MDM, and ADX

Most often, the three indicators, PDM, MDM, and ADX, are used to complement a trending strategy by improving the entry or exit timing. However, Colby[22] has found that using a 2-day calculation period and the following rules, the Dow Jones Industrial Averages were profitable for 72 years from 1928 to 2000:

Enter a new long position when the 2-day PDI > the 2-day MDI *or* the 2-day ADX > the 2-day smoothing of the ADX.

Close out a long position when the 2-day PDI < the 2-day MDI *or* the 2-day ADX < the 2-day smoothing of the ADX.

Entering a new short sale and exiting a short sale (buying to cover) are the opposite rules. All orders are executed on the close. The results show that a $100 investment would have returned $9,988 after profits were reinvested, which was better than a

---

[22]Robert W. Colby, *The Encyclopedia of Technical Market Indicators*, 2nd ed. (New York: McGraw-Hill, 2003).

TABLE 23.6 Profits or Losses from the Directional Movement Test, 2000–2011

Period	U.S. Bonds	DAX	EURUSD	Crude	Copper
4	21969	−106738	30688	28350	25800
8	−13500	−22063	21250	9990	29100
12	15250	−23500	58475	−11010	17275
16	10156	−36188	15775	−3880	72188
20	13813	−75563	50625	66400	23938
24	−37281	−5038	68425	54780	37288
28	−14688	−33000	90925	−5870	85525
32	−48250	19075	86150	−10310	107213
36	−55281	−8863	61000	48230	56525
40	−54531	−45413	56350	−12370	963
44	−43844	97775	59350	−20800	38863
48	−37438	79625	37925	−28960	103575
52	−42281	−2300	37175	−81440	99600
56	−23094	129625	14125	−102380	93425
60	−32719	130600	29738	−97750	116225
64	−33031	98588	19425	−104220	68075
68	−48313	88563	−12513	−132010	98300
72	−56938	71863	−56538	−118340	95488
76	−47281	66638	6938	−63060	44650
80	−46156	35688	11450	−34950	52325

buy-and-hold strategy by 118%. Because of the strong bull market of the 1990s, this strategy would have lost money on its short sales during that period.

Ranking Trends Using Prices

Chande also created his own trend ranking based on a rolling period of 50 or 100 days.[23] His *trend strength* is calculated as

$$S_t = \frac{\ln \dfrac{p_t}{p_{t-n}}}{\sigma\left(\ln \dfrac{p_t}{p_{t-1}}, n\right) \times \sqrt{n}}$$

Where $\sigma(\ )$ is the standard deviation of the 1-period returns over n periods. There is a similarity to the formula for noise given in Chapter 17. The numerator is the net return over n days, divided by a volatility function. In this case, S_t can be either positive or negative to show the trend direction.

---

[23]Tushar Chande, "Which Trends (and Markets) Are Best?" *Futures* (May 1997).

Kaufman's Strategy Selection Indicator

A market's *characteristic* is reflected in its price patterns. Some markets, such as the S&P 500 index, are very volatile with gradual upward moves and fast, sharp drops. In contrast, Eurodollar interest rates are very steady, often trading high volume at the same price. Qualifying markets by their underlying level of *noise* allows you to decide which trading strategy is most likely to be successful. Less noise favors trending systems and high noise makes mean-reverting techniques more appropriate.

The concept of *noise* has been discussed in Chapters 1 and 17, and it is again of great value when choosing which system to apply to a market and identifying when a market is best treated as trending. In essence, noise is market movement that has no direction or price movement in which the amount of direction is overwhelmed by erratic up and down movement. As a simple reminder, noise is defined using the *efficiency ratio*, *ER*, the absolute value of the net price change divided by the sum of the individual price changes taken as positive values, over the same time interval. Closing prices are used.

$$ER_t = \frac{|C_t - C_{t-n}|}{\sum_{i=t-n+1}^{t} |C_i - C_{i-1}|}$$

ER is not the same as volatility and indicates that noise increases as *ER* gets closer to zero because the divisor increases with the amount of noise. In Chapter 1, the level of noise was shown for world equity index markets (see Figure 1.8) and for Asian markets specifically. It was also shown that ranking the noise of various equity markets would show the maturity of that market (see "Maturing Markets and Globalization"). Emerging markets have less noise and more trend because there are fewer participants and those that are active seem to have the same view of the direction of the market, or the lack of liquidity drives prices further in one direction than would be seen in a highly liquid market. But equity index markets could be a special case. To see if this process can be generalized, we compare the returns from a 65-day simple moving average system with the 65-day (quarterly) average noise. A program that calculates the average efficiency ratio is *TSM Efficiency Ratio Average*, available on the Companion Website.

Results of Efficiency Ratio Selection

The results of calculating the average value of the efficiency ratio and the corresponding profit factor for a wide range of markets is shown in Figure 23.12 and Table 23.7. Data included 1990 through mid-2011, or whenever the data started. The left scale of Figure 23.12 shows the profit factor, the gross profits divided by the gross losses (a measure of return to risk) and along the bottom the efficiency ratio. Note that the range in the efficiency ratio is small because the calculation period of 65 days is long. It is not necessary to use more days than the longest daily run of upward or downward price moves. Including more days only compresses the values, but 65 days was used for consistency. Results show that there is a pattern going from the bottom left to the upper right, consistent with lower returns when the efficiency ratio is small (indicating more noise) and higher returns when the ratio moves toward 1.

Performance Versus Noise
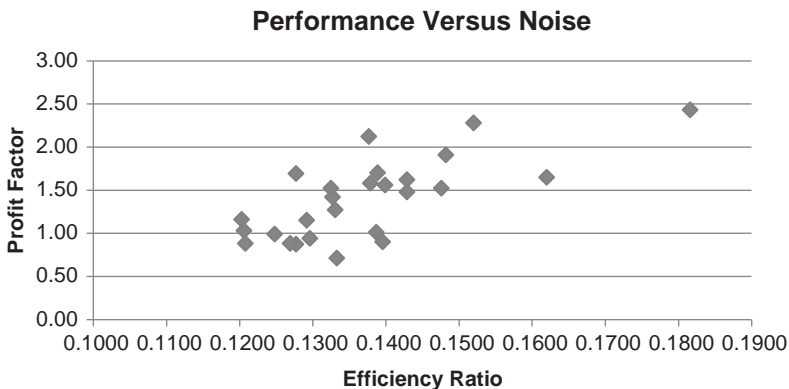
FIGURE 23.12 Scatter diagram of the profit factor versus the efficiency ratio for a wide range of futures and equities, 1990 through mid-2011.

TABLE 23.7 Values of the Efficiency Ratio (Noise) and Corresponding Profit Factor Sorted Highest to Lowest Average Noise

Market	Average Noise	Profit Factor
Eurodollar	0.1816	2.43
Sht Sterling	0.1620	1.65
AAPL	0.1520	2.28
Eurobund	0.1482	1.91
Nat gas	0.1476	1.52
Euro	0.1429	1.62
U.S. 10-year	0.1429	1.48
Japan	0.1399	1.56
IBM	0.1396	0.90
Copper	0.1389	1.70
AMZN	0.1387	1.01
DAX	0.1379	1.58
Crude	0.1377	2.12
MRK	0.1333	0.71
US bonds	0.1331	1.27
NASDAQ	0.1327	1.42
Cotton	0.1325	1.52
Gold	0.1296	0.94
Canada	0.1292	1.15
HO	0.1277	1.69
Soybeans	0.1277	0.87
MSFT	0.1269	0.88
Wheat	0.1248	0.99
GE	0.1208	0.88
CAC	0.1206	1.03
S&P	0.1203	1.16

Table 23.7 gives the numbers used in Figure 23.12. The highest efficiency ratio and largest profit factors are for the short-term interest rates and Apple stock, which has had a spectacular move during the past 10 years. The interest rates tend to have strong trends based on the way monetary policy is implemented, and the short maturities track the government rates more closely than the long maturities, which include a greater degree of speculation about future policy.

At the bottom of the table are most of the other stocks as well as the most active equity index markets and grains. In Chapter 1 it was shown that index markets have greater noise than other sectors, and the grains have less long-term direction because of seasonality and their response to changes in the U.S. dollar.

Figure 23.12 is consistent with the results in Chapters 1 and 17. Those markets with profit factors below 1.0 have net losses and are not good candidates for trend systems but may be best with mean-reversion strategies. Those farthest up and to the right are highly trending markets. When choosing markets to include in a diversified trading program, it is not necessary to include those that conflict with the nature of the strategy.

PROBABILITY OF SUCCESS AND RUIN

The relative size of trading profits and losses, the frequency of the losses, and the sequence in which they occur, combine to form an equity profile. This can apply equally to the returns of individual discretionary traders or fully automated trading systems. The profile can be used to determine the capitalization necessary to maintain trading during the losing periods and allow the system to continue uninterrupted trading. In investment terminology and probability theory, the level at which there is no longer enough money to continue trading is called the *point of ruin*, and the likelihood of getting there is the *risk of ruin*. The probability of the *risk of ruin* is expressed as

$$R = \left[\frac{1-A}{1+A} \right]^c$$

where $0 \leq R \leq 1$, and where 0 indicates no risk, and 1 is certain ruin

$A = P - (1 - P)$; P is the proportion of winning trades, also called the *trader's advantage*

c = the beginning units of trading capital (subsequent units can be expressed in fractional increments)

A trading system that has 60% profitable trades and trading capital in $10,000 units will have a risk of ruin calculated as follows:

$$A = 0.60 - (1 - .060) = 0.20$$

$$R = \left(\frac{1-0.20}{1+0.20} \right)^c = \left(\frac{0.80}{1.20} \right)^c = \left(\frac{1}{3} \right)^c$$
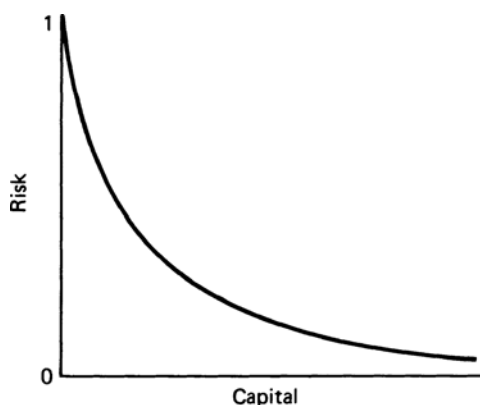
FIGURE 23.13 Risk of ruin based on invested capital.

When $c = 1$ unit, and the initial investment is $10,000, the risk of ruin $R = 0.33$, or 33%. When $c = 2$ units and the initial investment is $20,000, $R = 0.11$, or 11%. Therefore, the greater the trader's advantage or the greater the capital, the smaller the risk of ruin (Figure 23.13).

When using profit goals, the point at which trading would stop if the goal was achieved, the chance of ruin should decrease as the goal becomes closer. The relationship would be expressed as:

$$R = \frac{[(1+A)/(1-A)]^G - 1}{[(1+A)/(1-A)]^{c+G} - 1}, \ 0 \le R \le 1$$

where all terms are the same as above, and G is the goal in units of trading capital.

Wins Not Equal to Losses

The basic equations in the previous section are generally applied to gambling situations, where the size of profits and losses are the same. This requires that the percentage of winning events exceed the losing events in order to avoid ruin. Trading, however, often results in more losing trades than profitable ones and must therefore return much *larger* average profits than losses. This pattern is common to all conservation-of-capital systems, such as trend following. The risk of ruin of unequal profits and losses, including an unequal chance of a profit or loss, can be found as:[24]

C_T = the total capital available for trading (in units)
C_R = the cutoff point, where level of ruin is reached $(C_R < C_T)$
$C_A = C_T - C_R$, capital available to be risked

---

[24] Fred Gehm, *Quantitative Trading & Money Management*, rev. ed. (Chicago: Irwin, 1995).

E = the expected mean return per trade, the probability-weighted sum of values that a trade might take

$$E = \sum_{i=1}^{N}(PL_i \times p_i)$$

where PL_i = the possible profit or loss value
   p_i = the probability of PL_i occurring ($0 < p_i < 1$)
   N = the number of trades

E_2 is the expected squared mean return per trade, the probability-weighted sum of all the squared values of a trade,

$$E_2 = \sum_{i=1}^{N}(PL_i^2 \times p_i)$$

where PL_i and p_i are defined above.

$$D = C_A / \sqrt{E_2}$$
$$P = 0.5 + E / \left(2\sqrt{E_2}\right)$$

The risk of ruin is then

$$R = \left(\frac{1-P}{P}\right)^{D}$$

Introducing an objective and a desired level of capital L, the risk of ruin R becomes

$$R = 1 - \frac{[(1-P)/P]^D - 1}{[(1-P)/P]^G - 1}$$

where

$$G = 1 / \sqrt{E_2}$$

As in the first situation, using equal profits and losses, the risk increases as the objective L increases.

Ralph Vince[25] derived similar results from Griffin's work,[26] which claims to provide a "fair approximation" of risk. Vince's approach has been modified here for convenience and given in a way that allows a spreadsheet solution:

```
Risk of Ruin = ((1 - P)/P)^(MaxRisk/A)
```

[25] Ralph Vince, *Portfolio Management Formulas* (New York: John Wiley & Sons, 1990).

[26] Peter A. Griffin, *The Theory of Blackjack: The Compleat Card Counter's Guide to the Casino Game of 21*, 6th ed. (Las Vegas: Gamblers Press, 1999).

where the following terms are defined in the order needed for calculation:

AvgWin is the average winning trade (e.g., $400)
AvgLoss is the average losing trade (e.g., $200)
Investment is the amount invested (e.g., $10,000)
ProbWin is the probability (percentage) of a winning trade (e.g., 0.40)
ProbLoss is the probability (percentage) of a losing trade (e.g., 0.60)
MaxRisk is the maximum part of the investment that can be lost, in percent
    (e.g., 0.25)
AvgWin% is ABS(*AvgWin/Investment*)
AvgLoss% is ABS(*AvgLoss/Investment*)
Z is the sum of possible events, ($ProbWin \times AvgWin\%$) – ($ProbLoss \times AvgLoss\%$)
A is the square root of the sum of the squares of possible events, $((ProbWin \times$
    $AvgWin\%)^2 + (ProbLoss \times AvgLoss\%)^2)^{(1/2)}$
P is $0.5 \times (1 + (Z/A))$

A spreadsheet, *TSM Risk of Ruin*, and a function of the same name can be found on the Companion Website. Using the spreadsheet, Table 23.8 gives the risk of ruin for a few cases. Case 1 is the basic configuration using a $10,000 investment, a $400 average win, and a $200 average loss. The probability of a win or loss is 40% and 60% respectively, and the maximum loss in the account is set at 25%. That gives a risk of ruin less than 1%, shown at the bottom of the Case 1 column. In Cases 2 and 3, the size of the investment is halved, and the risk of ruin increases to 7.9% and 28.1%. The risk of ruin goes up faster than the decrease in the investment size. In Cases 4 and 5 the *AvgWin* is reduced from $400 to $350 and $325 and the risk of ruin also increases rapidly. In the last Case 5, the

TABLE 23.8 Examples of Risk of Ruin with Unequal Wins and Losses

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
AvgWin	400	400	400	350	325	325
AvgLoss	200	200	200	200	200	200
Investment	10000	5000	2500	10000	10000	10000
ProbWin	0.4	0.4	0.4	0.4	0.4	0.4
ProbLoss	0.6	0.6	0.6	0.6	0.6	0.6
MaxRisk	0.25	0.25	0.25	0.25	0.25	0.15
AvgWin%	0.04	0.08	0.16	0.035	0.0325	0.0325
AvgLoss%	0.02	0.04	0.08	0.02	0.02	0.02
Z	0.004	0.008	0.016	0.002	0.001	0.001
A	0.0200	0.0400	0.0800	0.0184	0.0177	0.0177
Z/A	0.2	0.2	0.2	0.1085	0.0565	0.0565
P	0.6	0.6	0.6	0.5542	0.5283	0.5283
Risk of Ruin	0.63%	7.93%	28.17%	5.22%	20.21%	38.31%

AvgWin is held at \$325 but the *MaxRisk* drops from 25% to 15% causing the risk of ruin to jump to 38.3%.[27]

ENTERING A POSITION

In the equities market, most people with retirement programs choose to add to their stock portfolio as money is available, and to reinvest profits. Over the long run, if the stock market continues to rise, that method is good. On the other hand, if a systematic trader gets an entry signal, they have only a few choices,

- Buy the entire amount at the time of the signal.
- Average into the position (or some weighted average).
- Wait for a better price.

We will compare these three choices, although the third, waiting for a better price, can be done at *any* better price or by applying some timing tool.

Averaging into a Position

It is not possible to cover all the different trading methods, but we can look at trend following as a popular example. We will take a small sample of markets, U.S. 30-year bonds, the *e*mini S&P, and gold, to represent a trending, noisy, and popular market. We apply an 80-day moving average from 1990 through mid-2011 as the benchmark case.

First, consider the results if entry points are randomly distributed. That is, if there is no significance in the point at which a trend changes direction, then over time the technique of averaging into a trade should give the same results as simply entering the entire position at the point of the initial trend change.

The test will look at always entering 5 contracts (or in equities, 5 shares), spaced from 1 to 5 days apart. That is, the first case will enter 1 contract each day for 5 days, the next case 1 contract every 2 days for 10 days, and so on. The slowest will average every 5 days over 21 days (days 1, 6, 11, 16, and 21). The maximum range of 21 days would be too long for a fast trend system, so it is likely that the number of days used to average into a trade will depend on the average holding period of the trend system. This study simply looks at one trend period, 80 days.

When averaging into a trade, if the trend changes direction before the full position is set, then the loss on the trade will be less than if you had entered all contracts on day 1. On the other hand, if prices were to move steadily higher after a new long signal, then the average entry price will be worse than day 1, the point where the trend first changed.

---

[27] A full discussion of risk of ruin can be found in Ralph Vince, *The Handbook of Portfolio Mathematics* (Hoboken, NJ: John Wiley & Sons, 2007), Chapter 12.

TABLE 23.9 Results of Averaging into a New Position Based on an 80-Day Moving Average System. The notation "5x2" heading the column means 5 total contracts entered 2 days apart.

		Total PL					
		MA80	**5 × 1**	**5 × 2**	**5 × 3**	**5 × 4**	**5 × 5**
U.S. Bonds	Volatility	53047	51559	50482	49487	48626	47765
	@15%	353646	343728	336548	329913	324173	318435
	Total PL	454375	440688	461719	429000	399844	358625
	Ratio	**1.28**	**1.28**	**1.37**	**1.30**	**1.23**	**1.13**
*e*mini S&P	Volatility	241349	10128	46535	45284	44193	43290
	@15%	1608997	67522	310231	301892	294618	288597
	Total PL	−212375	−35175	−2950	30775	40825	51362.5
	Ratio	**−0.13**	**−0.52**	**−0.01**	**0.10**	**0.14**	**0.18**
Gold	Volatility	60732	58617	57227	56184	55289	54546
	@15%	404879	390780	381516	374558	368593	363638
	Total PL	117000	161260	179890	173950	172240	166750
	Ratio	**0.29**	**0.41**	**0.47**	**0.46**	**0.47**	**0.46**

Table 23.9 shows the results. The first column of values, *MA80*, is the benchmark 80-day moving average where the entire position of 5 contracts is set at the time of initial entry. The next column gives the results of setting 1 contract on each of the first 5 days beginning on the day of the initial entry, The next column, 5×2, sets 1 contract on the initial day, the next four on days 3, 5, 7, and 9 (spaced every 2 days). The remaining columns space the entries every 3, 4, and 5 days, respectively. Then the final column takes 21 days to enter a trade, days 1, 6, 11, 16, and 21. The first row for each market is the annualized volatility of the daily profits and losses. The second row is the first row divided by 15%, which gives an estimate of the investment size needed if the target volatility of the portfolio is 15%. This allows a comparison across markets. The third row is the total profit from the 20-year test, and the fourth row is the ratio of the total profit divided by the investment (row 3). The ratio gives a simple way to compare all results.

Bonds are a trending market, so it would not be clear if waiting to enter is better than entering immediately; however, results improve to their best point by waiting 2 days between entries. With the *e*mini S&P, which represents a highly noisy market, results continue to improve as entries are spaced out. The waiting process turns results from negative to position. For gold, waiting is also an improvement but results seem to stabilize after 2 days. If we chose to wait 2 days between each of the entry orders, all three of the markets would show improvements.

It will be very difficult to generalize these results to all systems, even all trend systems, because the delayed entry must depend on the speed of the trend. If we had used a 10-day moving average instead of 80 days, then entering over 21 days does not make any sense. The total time to enter should be proportional to the speed of the trend or the average holding period of the trade. It is also not clear how a mean-reverting strategy

would fare by averaging in, but it might also improve. Averaging into a trade removes the uncertainty of a single entry and replaces it with a more stable value. In doing that, there seems to be an overall improvement in trend-following results.

A program, *TSM Averaging In*, available on the Companion Website, was used to produce the results.

Waiting for a Better Price

A simple way to improve an entry is to wait for a price that is better than the system signal. The obvious problem is when prices give a new signal and keep going and there is never a chance to enter at a better price. Statistically, it is possible that many small improvements will offset the occasional time when a trade is missed completely. To avoid the worst-case scenario of missing the biggest profit, we can enter at the close if no opportunity has occurred within 5 days, with the results shown in Table 23.10. If no opportunity occurred, the trade was entered on close of the fifth day. There is one explanation needed. The test platform used does not post the open position profit or loss on the last trade. For fast systems, where the trades have a short holding period, that makes little difference. For the 80-day moving average, which was the basis of this test, the last trade may have lasted for many months, as it did with gold. Had the last trade been included, as it was in the previous study, gold would show a large net profit. That does not change the effect of this rule, but should explain any inconsistencies between the results of the benchmark case here and in the previous example. The program, *TSM Wait for a Better Entry*, is available on the Companion Website. It allows you to change both the threshold and the maximum window in which the threshold will be applied.

We see that waiting, even for a minimum move pullback after the initial signal, results in a significant improvement in performance for the three markets during the period from 1990 through mid-2011. However, a different range of thresholds was used for each market because the threshold is defined as a point value. It may be worth trying this with a percentage, but back-adjusted futures prices can be distorted when the data goes back many years, so you need to be careful about using percentages.

The minimum threshold, shown in the second row of values, cut the number of trades by about 40% in all three markets. This shows that there are many trend trades that lasted only a few days, and, because of waiting for a better price, the trend turned back in the previous direction, and the false signal was not taken. The minimum threshold also turned the gold results from a loss to a profit, while it cut the S&P losses in half. It had the least improvement in bonds, which has the strongest trend. We would expect markets with a large noise component to work well with this rule, so that the S&P had the biggest improvement even though it never showed a profit. A secondary benefit is that the profits per trade improved significantly.

The need for a maximum number of days in which this rules is applied is a critical element for its success. While these results show improvements, it is possible that without the 5-day window, one exceptionally profitable trade would be missed. Even though it does not show in the tests, the possibility of that problem makes using a window necessary.

TABLE 23.10 Results Based on Waiting for a Better Price after the Initial Entry Signal

U.S. 30-Year Bonds					emini S&P					Gold				
Threshold	Total PL	Trades	% Prof TR	Prof /TR	Threshold	Total PL	Trades	% Prof TR	Prof / TR	Threshold	Total PL	Trades	% Prof TR	Prof/ TR
None	61468	188	36.1	327	None	-44162	263	30.4	-168	None	-1580	246	32.9	-6
Min	72781	113	41.5	644	Min	-22337	155	26.4	-144	Min	12170	147	37.4	83
0.03125	71469	112	42.0	638	0.5	-22575	155	27.7	-146	1	10990	130	33.8	85
0.06250	69656	110	42.7	633	1.0	-21850	153	28.1	-143	2	14630	121	33.1	121
0.09375	69469	108	44.4	643	1.5	-22475	151	27.2	-149	3	14990	116	32.8	129
0.12500	69594	108	44.4	644	2.0	-20588	149	28.9	-138	4	14260	112	31.3	127
0.15625	68219	108	42.6	632	2.5	-22063	148	29.1	-149	5	15010	112	31.3	134
0.18750	69188	102	43.1	678	3.0	-20438	146	29.5	-140	6	15070	112	31.3	135
0.21875	70156	101	43.6	695	3.5	-18163	143	30.1	-127	7	14910	111	31.5	134
0.25000	73281	100	46.0	733	4.0	-14575	140	30.0	-104	8	15930	110	31.8	145
0.28125	71313	100	46.0	713	4.5	-14238	140	30.7	-102	9	16600	109	32.1	152
0.31250	71406	99	45.5	721	5.0	-13888	139	30.9	-100	10	12560	109	33.0	115
0.34375	72594	98	45.9	741	5.5	-11588	137	31.4	-85	11	13700	108	33.3	127
0.37500	70750	97	46.4	729	6.0	-11713	136	31.6	-86	12	13700	108	33.3	127
0.40625	67938	97	45.4	700	6.5	-11713	136	31.6	-86	13	13700	108	33.3	127
0.43750	69938	94	46.8	744	7.0	-11550	135	31.9	-86	14	14800	107	33.6	138
0.46875	69313	93	47.3	745	7.5	-10125	134	32.1	-76	15	15810	107	33.6	148
0.50000	70594	93	48.4	759	8.0	-9275	134	32.1	-69	16	19000	105	34.3	181
					8.5	-9175	134	32.1	-68	17	20890	103	34.0	203
					9.0	-9588	133	32.3	-72	18	20890	103	34.0	203
					9.5	-10250	131	32.1	-78	19	22420	103	34.0	218
					10.0	-9700	131	32.1	-74	20	22420	103	34.0	218

TABLE 23.11 Results of Timing the Entry Using an 8-Day RSI and an 80-Day Moving Average, and Threshold Entry Levels of 30 and 70 for Buy and Sell Signals over 20 Years

Market	Total PL	Trades	% Prof Trades	Profit/ Trade
Bonds	65750	90	47.7	731
S&P	−10550	124	34.6	−85
Gold	22180	105	34.2	211

Using Timing for a Better Price

The last study applies a momentum indicator to find one better entry point after the initial trend signal. Linda Raschke's *First Cross* system used this as a primary rule for short-term trading. In this case, we will continue to use the 80-day moving average as a benchmark, and the popular RSI as a timing indicator. Long positions will be entered after the initial signal, or at the time of the initial signal, if the RSI is below a threshold level set at 30. The threshold for selling was 70. Again, a maximum window will be used to avoid missing a trade. The RSI will be 8 days rather than the default 14 days in order to find more opportunities within the window.

Table 23.11 shows results similar to the previous tests but with fewer trades. Even an 8-day RSI does not generate as many timing signals as the other studies, so that the trends reverse more often within the 5-day window without an entry occurring. The program, *TSM Timing the Entry*, available on the Companion Website, was used to produce these results.

From the three studies, we can conclude that these three markets, and hopefully markets in general, benefit from entry timing. These techniques were very specific, using a long-term trend as a benchmark, but the same concept should apply across a reasonably broad set of trend calculation periods until we get to short-term trading. When trades are held for only a few days, the application may require using hourly data to time the entry instead of daily data. Still, prices do not go only one way, and market noise should allow us a better place to enter a trade.

COMPOUNDING A POSITION

At some point, most traders find themselves adding to, or compounding, their position (or seriously considering it). Some view this as a means of concentrating their resources on those markets that have more potential. There are two lines of thinking. When a trade becomes more profitable, it is confirming its move and is thought to deserve more of a commitment than a trade that has not become profitable. The alternative, when prices move against you, is *scaled-down buying*, or *averaging down*, popular in the securities industry. In pension programs, where new money is added at regular intervals, regardless of market movement, the timing is arbitrary, and profits are entirely dependent on whether the stock market as a whole, or the fund in which you are invested, is higher

in the long run. But that is not trading. If positions are not added specifically based on profits, losses, or some timing method, then it is the same as averaging into a position.

The types of compounding are illustrated in Figure 23.14. Part a shows smaller positions added, with the added size on the left and the total position on the right. Part b shows equal positions added, and part c shows adding then removing positions. Systematic traders may use a number of conditions for entering addition size, ideally,

- A sustained basing formation
- Expectation of substantial additional profits
- A clear risk level

They expect profits to accumulate soon after adding positions
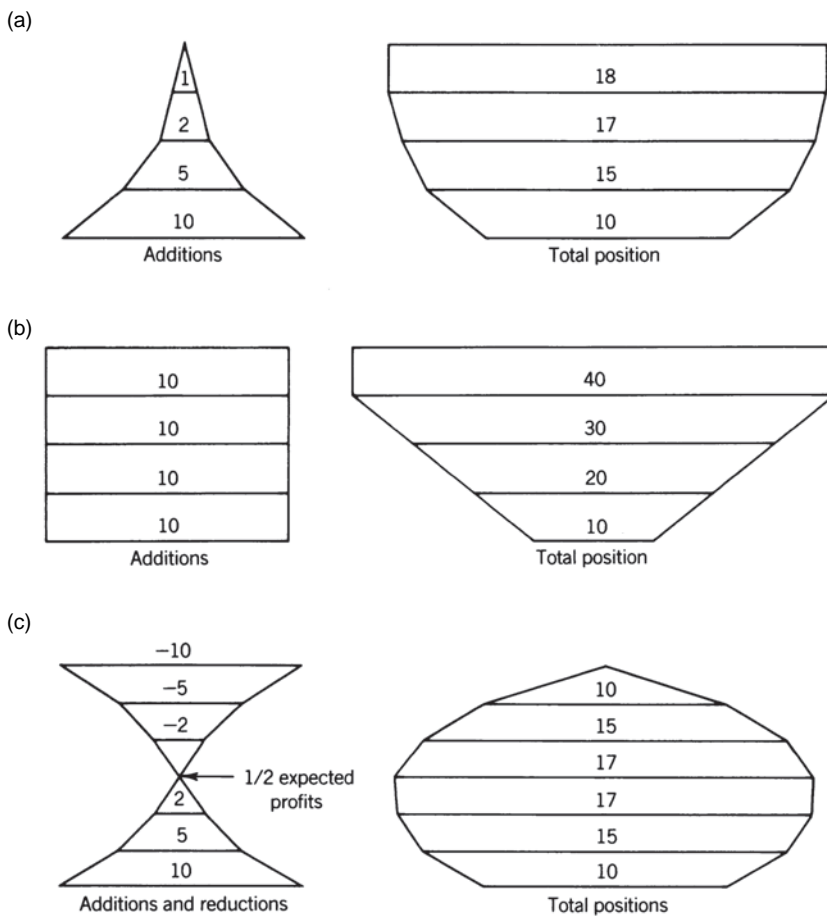
FIGURE 23.14 Compounding structures. (a) Scaled-down size (upright pyramid) offers a small amount of compounding. (b) Adding equal amounts (inverted pyramid) gives maximum leverage. (c) Reflecting pyramid combines leverage and profit-taking.

No matter how well each entry is chosen, compounding will result in the largest holdings at the highest (or lowest) price; when the market reverses, losses occur on a larger base, and profits can disappear quickly. A compounded position can be very fragile and top-heavy; prices must be watched carefully for a change of direction. If you continue to leverage up with all available funds, then risk will become unmanageable.

Now consider building a position, similar to averaging in, rather than leveraging. If you plan to buy 1000 shares of Apple or 10 contracts of S&P, then you can start with 200 shares or 2 contracts and add the same amount at regular intervals, but only if you have new high profits. If prices go the wrong way, then you wait. If the trend reverses, the loss will be smaller if the maximum position was not set. Results could be similar to the previous study on averaging into a trade.

Adding on profits is not the same as waiting for enough profit and using those profits to add size. With both stocks and futures, the amount of profit would need to be very large to have enough to add, so the opportunities are small. For this method, there must be adequate funds set aside to carry the full position.

In the rest of this section, we will look at a few limited cases of entering a position in parts based on market conditions:

- Adding equal amounts on new high profits with a minimum number of days separating the entries
- Adding half the previous size with the same conditions as above
- Adding equal amounts on new losses with a minimum number of days between entries.

Intuitively, the last method should give a better entry price but might miss the move. In addition, if trend-following works, then we want the entire position set during those trades that capture the fat tail.

Adding on Profits

As with the previous studies, it is necessary to choose a trading system, and again we will use a simple moving average, but with a calculation period of 60 days. We will look at four markets, Apple (AAPL), Amazon (AMZN), 10-year Treasury notes (TY), and the *e*mini S&P (ES). Both Apple and Amazon are popular stocks and have made big moves, 10-year notes are very trending, and the S&P is noisy, which gives us a somewhat varied sample. The data covers the past 20 years. The rules will be

1. Enter 1 unit on the new trend signal.
2. Enter another unit after n days only if the trade is on new high profits.
3. Stop entering after five units have been added.

Looking first at Apple, the first entry in Table 23.12, the largest total profits occur with the simple 60-day moving average without additional entries; however, the *Profit*

TABLE 23.12 Building a Position on New High Profits

	Method	Total PL	Prof Factor	Trades	Profits/ Trade	Units Entered
Apple	MA 60	430.35	2.79	238	1.810	5.00
	1 Day	389.34	3.42	624	0.624	2.62
	3 Days	374.54	4.18	501	0.748	2.11
	5 Days	350.82	4.34	473	0.742	1.99
	10 Days	305.34	4.33	425	0.718	1.79
Amazon	MA 60	153.08	1.45	172	0.890	5.00
	1 Day	96.98	1.43	425	0.228	2.47
	3 Days	76.90	1.36	389	0.198	2.26
	5 Days	71.74	1.36	348	0.206	2.02
	10 Days	56.72	1.35	307	0.185	1.78
Notes	MA 60	13734	1.11	292	47.04	5.00
	1 Day	29825	1.42	647	46.08	2.22
	3 Days	30791	1.49	569	54.10	1.95
	5 Days	27700	1.48	519	53.36	1.78
	10 Days	17828	1.34	464	38.42	1.59
S&P	MA 60	22162	1.14	301	73.00	5.00
	1 Day	28875	1.28	719	40.16	2.39
	3 Days	28225	1.3	626	45.08	2.08
	5 Days	34463	1.43	564	61.10	1.87
	10 Days	30728	1.46	484	63.48	1.61

Factor (gross profits divided by gross losses) improves steadily as the gap between entries increases. At the same time, the average return per share (*Profits/Trade*) will decline because the average entry price will increase for longs and decrease for short sales. On the far right, the *Units Entered* give us an idea of how many units were traded on average. For example, for the 1-day wait, 2.62 units indicates that the average trade had 2.62 shares while the 60-day moving average would always have 5 shares. Amazon, the second set, is the exception, showing a decline in the profit factor.

The two futures markets both show improved total profits as well as higher profit factors as trades are entered over time; 10-year notes peak with a 3-day wait while the *e*mini S&P peaks at 5 days for profits and 10 days for the profit factor. Overall, entering as profits accumulate rather than entering all at one time gives much better results for the two futures markets and reasonably good results for the two stocks. The program, *TSM Adding on Profits*, used to create Table 23.12, can be found on the Companion Website.

There are many combinations that could have been tested, a wide range of trend speeds, a mean reversion system, and different ways of adding to the trade. Analysts will always need to study these techniques in a conceptual framework as well as how it applies specifically to their set of trading systems.

Averaging Down

Because the practice of averaging down is pervasive in the equities markets, we would be remiss if we did not run the same test as *Adding on Profits* but reverse the logic so that we are *adding on losses*. Equal positions were added after at least n days provided that the open position was making new losses greater than remaining at the previous loss. There was an additional weaker test that allowed new positions to be entered as long as the trader showed a loss every n days, but those results were actually worse than this study.

Table 23.13 shows that Apple was the only market in which averaging down improved long-term performance, both in total profits and profit factor. The other three test markets produced noticeably worse results for all combinations of delay, given the same trend period of 60 days. Of course, other trend periods might have given different results, but a 60-day moving average is representative of the faster end of the macrotrend space and generally a conservative choice. Although a brief test, we would expect that any of the other methods, averaging in or adding on profits, are likely to yield better results. Readers interested in testing this will find the program *TSM Averaging Down* on the Companion Website.

TABLE 23.13 Averaging Down by Adding Equal Amounts Whenever There Was a New Low Separated by a Minimum Number of Days

	Method	Total PL	Prof Factor	Trades	Profits/ Trade	Units Entered
Apple	MA 60	430.35	2.79	238	1.810	
	1 Day	877.4	3.36	377	2.330	1.58
	3 Days	672.07	3.32	308	2.180	1.29
	5 Days	505.64	2.81	284	1.780	1.19
	10 Days	502.1	2.91	263	1.910	1.11
Amazon	MA 60	153.08	1.45	172	0.890	
	1 Day	−19.55	0.97	278	−0.070	1.62
	3 Days	65.83	1.13	226	0.290	1.31
	5 Days	70.13	1.16	207	0.340	1.20
	10 Days	102.42	1.26	191	0.540	1.11
Notes	MA 60	13734	1.11	292	47.04	
	1 Day	−7234	0.97	455	−15.90	1.56
	3 Days	7171	1.04	377	19.02	1.29
	5 Days	7000	1.04	357	19.61	1.22
	10 Days	8015	1.05	323	24.82	1.11
S&P	MA 60	22162	1.14	301	73.00	
	1 Day	−2475	0.99	456	−5.43	1.51
	3 Days	5437	1.03	378	14.38	1.26
	5 Days	10737	1.06	351	30.59	1.17
	10 Days	23712	1.14	324	73.19	1.08

EQUITY TRENDS

Every system has profit and loss patterns that can be seen clearly by plotting its daily or weekly returns. Long-term trending techniques show that once or twice each year, there are major increases in profits corresponding to a trending market; at other times, returns show a steady decline or a stabilizing pattern.

Increasing position size as equity increases during a trending period always results in being fully invested at the top of the cycle, when losses begin. Losses will be on a larger base than profits, and equity will drop much faster than it increased. Decreasing size as profits increase will reduce risk but lose the advantage of capturing the fat tail. Similar changes can occur at the bottom of an equity cycle after a nontrending market period. A sustained losing streak may cause a speculator to reduce the investment in proportion to dwindling capital. If this happens, the result will be entering into a profitable period with a smaller investment than the prior losing period. The system must have disproportionately larger profits to recover the losses and achieve a net gain.

A classic example of a pattern in returns is one where there is a 100% gain for each profitable period, followed by an interval where there is a 50% loss. If this cycle is repeated twice each year then each year will have the following equity pattern:

	Change in Equity	Total Equity	
Original Investment		$10,000	
Gain of 100%	+10,000	20,000	First 6 months
Loss of 50%	−10,000	10,000	
Gain of 100%	+10,000	20,000	Second 6 months
Loss of 50%	−10,000	10,000	

Trading would be a great deal of effort for no return.

Some investors choose not to compound their returns, but let profits accumulate as *reserves*, also called *account excess*. Given the same profit/loss cycle as above, but holding the investment constant results in a very different equity pattern. The management of the reserves is the key to successful management. Starting with margin and reserves equal, reserves increase during profitable periods and decrease during losing ones. Proportionately more of the total equity is traded during losing phases. The results are:

	Change in Equity	Margin	Reserves	Total Equity	Reserves/ Equity
Original investment		10,000	10,000	20,000	50%
Gain of 100%	+10,000	10,000	20,000	30,000	67%
Loss of 50%	−5,000	10,000	15,000	25,000	60%
Gain of 100%	+10,000	10,000	25,000	35,000	71%
Loss of 50%	-5,000	10,000	20,000	30,000	67%

Using the natural pattern of trading returns, hold the number of positions the same and allow the reserve to increase during profitable periods; maintain the same position size through the beginning of the next losing period. When the drawdown has slowed or stabilized, the total account value can be redistributed into margin and excess according to the original 50% formula. In the next example, the total account value of $25,000 is distributed 40% to margin and 60% to reserve at the end of the first cycle. It is redistributed so that the next profit phase will be entered with a larger base than the previous losing cycle. The result is a gradual increase in profits:

	Change in Equity	Margin	Reserves	Total Equity	Reserves/ Equity
Original investment		10,000	10,000	20,000	50%
Gain of 100%	+ 10,000	10,000	20,000	30,000	67%
Loss of 50%	−5,000	10,000	15,000	25,000	60%
Redistribute		12,500	12,500	25,000	50%
Gain of 100%	+10,000	12,500	25,000	37,500	67%
Loss of 50%	−5,000	12,500	18,750	31,250	60%
Redistribute		15,625	15,625	31,250	50%

Trading on Equity Trends

It seems reasonable to think that the equity stream generated from a trend-following system, especially a long-term macrotrend system, would also trend. We would expect prices to move in a favorable direction the majority of the time, producing a chart with reasonably smooth swings. By applying a simple moving average to the accumulated profits and losses, we could exit positions when the trend of equity is down.

Using the results from an 80-day moving average applied to the same futures markets used earlier, 30-year bonds, the *e*mini S&P, and gold, we applied both a 10-day and 60-day moving average to the cumulative profits over 20 years. When the trend of the equity was up on day t compared to $t - 1$, we used the returns on day $t + 1$; otherwise, we used a return of zero. While this is only one calculation period, 80 days represents the heart of the macrotrend system. For simplicity, only 1 contract was traded on each signal rather than a risk-adjusted position size. The results are shown in Figures 23.15a–c and the calculations in the spreadsheet *TSM Equity Trends* are available on the Companion Website.

None of the three cases show consistency, even though there were improvements for the S&P and gold using both the 10-day and 60-day equity trends. Bonds, the market with the most trend, show a significant deterioration using this method. The equity trend method also did not include any switching costs, which would be more significant for the 10-day equity trend. We can infer that this method would perform poorly on a diversified portfolio, which should have shorter periods of negative movement. Of course, this does
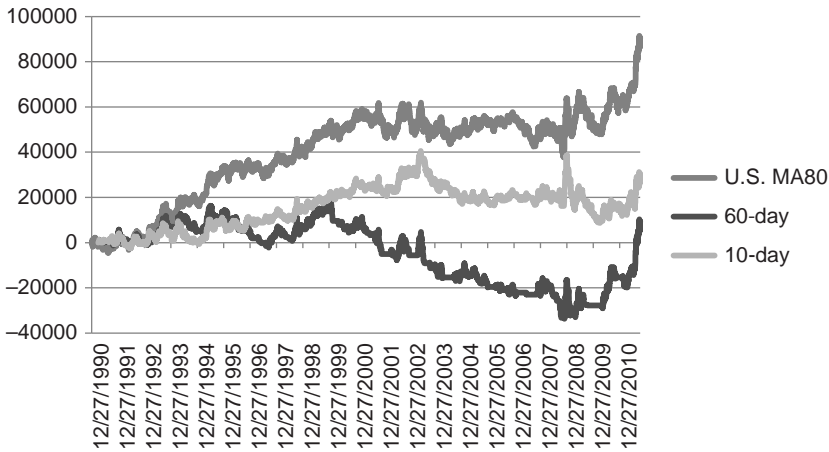
FIGURE 23.15a   30-year bonds cumulative profits shown with a 10-day
and 60-day equity trend calculation.
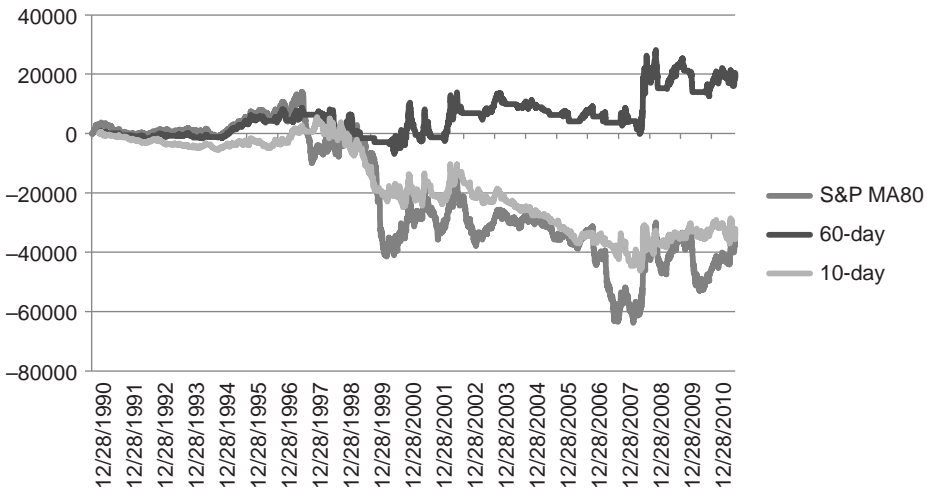
FIGURE 23.15b   The *e*mini S&P cumulative profits shown with a 10-day and 60-day
equity trend calculation.

not mean that a solution cannot be found, only that a simple approach to finding trends
in systematic equity patterns is not going to be easy.

In pursuing this, you should first find out if the returns are positively autocorrelated,
meaning that there is persistence in either positive or negative returns. You must also
consider that applying this to hypothetical equity patterns is not likely to work because
hypothetical results are always better than actual trading results.
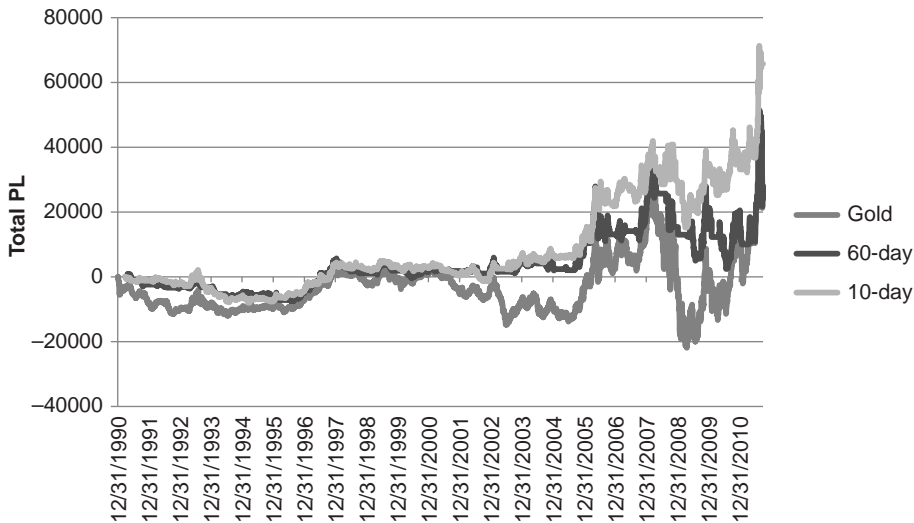
FIGURE 23.15c   Gold cumulative profits shown with a 10-day and 60-day equity
trend calculation.

INVESTING AND REINVESTING: OPTIMAL *f*

Optimal f is the optimal fixed fraction of an account that should be invested at any one
time, or the size of the bet to place on any one trade. The amount to be risked is measured
as a percentage of the portfolio size. The objective is to maximize the amount invested
(put at risk) yet avoid the possibility of a loss greater than some target amount. Trading a
very small part of assets can be a poor use of capital, while trading too much guarantees
bankruptcy or ruin. Optimal *f* is the ideal portion of an investment that should be placed
at risk at any one time.

　　Investing generally has a 2-level optimal *f:* (1) the part of the total portfolio put at
risk compared to that part held in cash equivalents, and (2) the individual size of the
commitment to each stock or futures contract within that portfolio. This is particularly
important for futures, where the high leverage of individual markets makes it very easy
to risk too much on each trade.

Risk Assessment

Then, how much should be invested? The optimal amount is difficult to pinpoint because
you would have to know what risks lie ahead, and, of course, that's not possible.

　　The most extreme situation can be found by using the calculation for *risk of ruin* in
a previous section of this chapter, most often applied to gambling situations in which the
bet sizes, payout, and odds are well defined. When there are enough test data and trades,
this technique has been applied to trading systems (see the section "Wins Not Equal to

Losses"). All this depends on using historic data, which is often not a reliable indication of future price movement, which seems to get more correlated and more volatile.

Some analysts have tried to deal with the uncertainties of price movements by using a *Monte Carlo* technique in testing, which moves blocks of data, or sequences of returns, using a random process. This results in a distribution of possible risk combinations from which you can assign the probability of loss. Monte Carlo risk analysis is considered unusually severe because many of the worst combinations happen because the return series is identified without regard to possible cause-and-effect relationships. For example, a long-term trend-following technique expects to capture moves that are based on economic or government policy; therefore, sustained profits are most often followed by a reversal before the trade is ended. In fact, the size of the change of direction at the end of a trend is directly related to magnitude of the prior trend because of the lag. To move this data around so that the loss at the end of the trend comes at a different time may create a large loss without the preceding profit, a situation that is unfair to the trading strategy. Yet even an analysis of real performance is likely to understate the size of the future risk.

For an initial investment, optimal *f* is simply the maximum part of that portfolio that can safely be traded without any significant risk of ruin. For those investors who withdraw profits as they occur and continue to trade based on the same assumed initial investment, nothing need be changed unless exceptionally high risk requires a decrease in the amount of leverage. However, it is more common—and more complicated—for the investor to vary the amount committed to the market by either increasing or decreasing leverage. This involves (1) determining the right time to change the leverage, (2) calculating the amount to increase the investment when there are profits, and (3) figuring the size of the reduction when there are losses exceeding some designated amount. These are issues that are addressed by optimal *f*.

Finding Optimal *f*

Ralph Vince, in his popular book *Portfolio Management Formulas*,[28] focuses on optimal *f*, risk of ruin, and other practical items. The significance of this approach is the need to maximize the amount invested yet avoid the risk of ruin. Optimal *f* is the ideal amount of an investment that should be put at risk at any one time. First, we need to know what percentage gain is required to recover a percentage loss.

$$Required\ gain = \frac{1}{1 - Percent\ loss} - 1$$

That is, a 50% loss requires a 100% gain to restore the original value. Because the amount risked on each trade depends on our expectations of loss, the results obtained

---

[28] Ralph Vince, *Portfolio Management Formulas* (New York: John Wiley & Sons, 1990), 79–86. Also see Ralph Vince, *The Handbook of Portfolio Mathematics* (Hoboken, NJ: John Wiley & Sons, 2007), Chapters 4 and 5.

from the optimal f calculation will be the size of the bet, the invested amount, or the number of futures contracts to be traded, as a percentage of the maximum loss. The value used as a maximum loss will be an estimate, because losses can always be greater than those already experienced in the market, or those found by historic testing. In addition, the optimal f will be different for each system, depending upon its performance profile.

The mathematics needed to determine optimal f is based on the Kelly Betting System.[29] Kelly states that the optimum bet is the one that maximizes the growth function $G(f)$:

$$G(f) = P \times \ln(1 + B \times f) + (1 - P) \times \ln(1 - f)$$

where f = the optimum fixed fraction
         P = the probability of a winning bet or trade
         B = the ratio of the average winning return to the average losing return
         ln = the natural log function

The solution for finding the optimal fixed fraction to invest uses the geometric product and geometric mean, which represent the way in which profits and losses accrue.

$$Optimal\ f = \max\left(\prod_{i=1}^{n}\left(1 + \frac{f \times (-R_i)}{Largest\ loss}\right)\right)^{1/n}, \quad \text{for } f = 0.01 \text{ to } 1.0$$

where   max = the function that returns the maximum value
          Π = the product function
         R_i = the series of individual trade returns
          n = the number of trades

By testing values of f between 0.01 and 1.0, and finding the geometric mean of all trades (each percentage profit or loss applied to the account value before the current trade), the value of f is found that gives the best return. That f-value is the optimal f, the percentage of the total account that should be invested for each trade. Another way of expressing optimal f, is given as[30]

$$f = [p \times (PLR + 1) - 1] / PLR$$

---

[29] John L. Kelly, Jr., "Kelly Betting System," *Bell System Technical Journal* (July 1956). Kelly's own method began with the optimal bet size based on unequal payouts as $p - (1 - p)/r$, where p is the probability of winning and r is the win/loss payout ratio.

[30] Robert P. Rotella, *The Elements of Successful Trading* (New York: The New York Institute of Finance, 1992), 549–550.

where   *PLR* = the ratio of average profit to average loss
             p = the probability of a winning trade

Therefore, if $p = 0.50$, there is an equal chance of a profit or a loss, and the average profit is \$400 while the average loss is \$200 (giving *PLR* = 2.0), then $f = (0.50(2 + 1) -1)/2 = 0.5/2 = 0.25$ or 25% of the available capital. Given an equal chance of a profit or a loss, it is not likely that there would be 4 losses in a row, each of 25%; however, the theory of runs shows that, out of every 100 trades, there should be one run of 6. Eventually, there will be a run of 4 or 5 losses in a row. Optimal *f*, however, invests a fraction of the current equity; therefore, after a loss of 25%, the next investment is 25% of the balance, or 18.75% of the initial equity. If there are further losses, that amount drops to 14.06%. After three losses in a row, instead of having lost 75% of the initial equity, the investment has only dropped by 57.81%. Over time, with profits twice as large as losses, and winning trades alternating normally with losing trades, the losses will be recovered.

Observations of Optimal *f*

According to Alex Elder,[31] there are some difficulties in using optimal *f*. Because the value is based on every historic trade, the ideal amount to invest on the next trade will keep changing. In addition, if you trade a position larger than determined by optimal *f*, and get average results, you can expect to go broke eventually because you are overinvesting. On the other hand, if you invest less than the optimal amount, then your risk decreases arithmetically, but your profits decrease geometrically, which is another bad scenario. Because this is too complicated for most investors, the simple solution is to keep trading the same amount, with a reserve sufficiently large to absorb most extreme, adverse price moves.

On the positive side, Dr. Elder concludes that the most useful result of optimal *f* is that it shows the trader to:

- Never average down.
- Never meet margin calls.
- Liquidate the worst position first.

The application of optimal *f* uses the end value of individual trades, somewhat similar to the payout of a bet, while trades are comprised of many days of individual returns. The drawdown during a trade could be much larger than the final outcome of that trade. Then there is more risk than can be seen using only the final trader result.

Markowitz also took a look at this problem in 1959, with the following approximation:

Expected log return = Expected return – ½ Variance of returns

---

[31] Dr. Alexander Elder, *Trading for a Living* (New York: John Wiley & Sons, 1995).

Using an amount of leverage M, the log returns becomes:

Expected leveraged log returns = M × Expected return – ½ M² Variance of returns

then

Optimal leverage = Returns/Variance = μ/σ²

As with other methods, this depends on the future being the same as the past. Because of the uncertainty, it may be best to consider using these methods only to reduce leverage, not increase it. Traders should measure success in terms of a return-to-risk ratio; therefore, this method of leveraging should be compared to other volatility stabilization approaches given in the next chapter.

COMPARING EXPECTED AND ACTUAL RESULTS

In the development of an economic model or trading system, the final selection, as well as the choices made along the way, are based on comparing the results of one method with another. Often the results are given in terms of information ratios, profit/loss ratios, annualized percentage profits, expected reliability, and drawdown ratios. Although these statistics are common, their predictive qualities and sometimes their accuracy are not known. On occasion, these results are generated by a sample that is too small; usually they are not the results of actual performance but an historic test. This does not mean that the model will be unsuccessful, but that the pattern of success might vary far from expectations. In actual trading, everyone experiences a series of losses far exceeding the maximum level that was expected; at that point, it is best to know whether this situation could occur within the realm of the system's profile or whether the system has failed. For example, a moving-average system is expected to have about 33% profitable trades with an average profit to average loss ratio of 4:1. But the first 10 trades of the system are losers. Should trading be stopped?[32]

Binomial Probability

Consider the application of a random-number sequence to the trading model. What is the probability of l losses in n trades when the probability of a loss is p? Most of the work in this area of probability is credited to Bernoulli, whose study of a random walk is called a *Bernoulli process*. A clear representation of a random walk is shown by Pascal's triangle (Figure 23.16), where each box represents the probability of being in a particular position at a specific time in a forward random walk. The result of this process is called a *binomial distribution.*

---

[32] Another technique, the probability of a drawdown was discussed in Chapter 21, in the section "Is the Model Broken?"
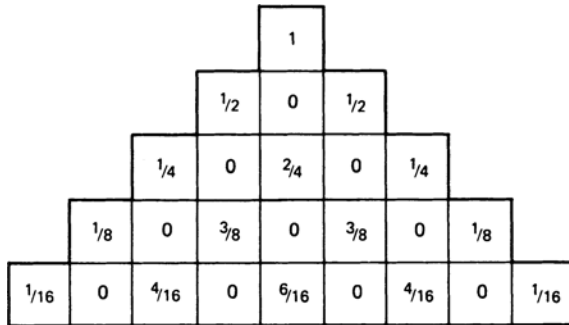
FIGURE 23.16 Pascal's triangle.

The forward random walk has an analogy to price movement, with the far edges of Pascal's triangle showing the probability of a continuous sequence of wins or losses using random numbers. The sequence $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots, (\frac{1}{2})^n$ is exactly the same as in the discussion of the Theory of Runs. The probability of successive losses can be calculated as the likelihood of a run of the same length, $(\frac{1}{2})^{n+2}$.

A binomial distribution is useful in considering the total number of losses that can occur in any order within a sequence of trades; it is the probability of getting to a specific point at the base of Pascal's triangle when there is a high probability of moving to the left (losses) rather than the right (profits). The formula for the binomial probability is:

$$B(l:p,n) = \frac{n!}{l!(n-1)!} p^l (1-p)^{n-1}$$

where l = the number of losses
   n = the total number of tries
   p = the probability of a loss

and the symbol "!" is the *factorial* (e.g., $5! = 5 \times 4 \times 3 \times 2 \times 1$).

Consider the first 5 trades of a system with a probability of success of 1/3. How many losses should be expected? To answer the question, it must be phrased differently. In this case, what is the probability of having 4 losses out of the first 5 trades? Let $l = 4$ and find the binomial probability B for all possibilities based on a normal distribution function. Then,

$$B(4:0.667,5)\frac{5!}{4!1!} \times (0.667)^4 \times (0.333)^1$$

$$= \frac{120}{24} \times (0.19792) \times (0.333)$$

$$= 5 \times 0.659 = 0.32954$$

The binomial probability of having 4 losses out of the first 5 trades is about 33%. Table 23.14 shows the probability of loss for the first 5, 10, and 15 trades of a system with a ⅓ predicted reliability. Results show the highest probability of loss is at the ⅔ point

TABLE 23.14   The Probability of a Specific Number of Losses

5 Trades		10 Trades		15 Trades	
Losses	Probability (%)	Losses	Probability (%)	Losses	Probability (%)
0	1	0	0	0	0
1	4	1	0	1	0
2	16	2	0	2	0
3	33	3	2	3	0
4	33	4	5	4	0
5	13	5	14	5	1
		6	23	6	2
		7	26	7	6
		8	19	8	11
		9	9	9	18
		10	2	10	21
				11	20
				12	13
				13	6
				14	2
				15	0
m = 3⅓		m = 6⅔		m = 10	
sd = 1.05		sd = 1.5		sd = 1.825	

m = Mean.
sd = Standard deviation.

(mean) for each sequence, but the standard deviation gives the range of variance about the mean, so that from 2.3 to 4.4 losses are expected in every 5 trades, 5.2 to 8.2 in 10 trades, and 8.2 to 11.8 losses in 15 trades.

Note that in the 5-trade example, the chance of no loss is only 1% and there is a 13% chance of all losses. For the purpose of evaluation, it is easier to look at the maximum rather than the minimum number of losses. For 15 trades, there is an 8% chance of 13 or more losses; if the system has produced more than 12 losses in that period, there may be something wrong with the trading method.

Instead of the Pascal distribution, which is normal, the Poisson and various skewed distribution functions will be more appropriate for trading performance. It is well known that price changes and their returns have a skewed distribution with a fat tail.

χ^2—Chi-Square Test

Once a system has been traded and there is enough data to give a performance profile, a simple correlation between these actual results and the expected results can be found using the *chi-square test*. The sample error, $1/\sqrt{N}$, will tell you if there is enough data.

Assume that the real trading results show a reliability of 20% (1 out of 5 profitable trades) as compared to the expected reliability of 35%. What are the chances of getting the expected results? The chi-square test is

$$\chi^2 = \frac{(O-E)^2}{E}$$

where O = the observed or actual result
        E = the expected or theoretical result.

then,

$$\chi^2 = \frac{(20-35)^2}{35} + \frac{(80-65)^2}{65}$$
$$= \frac{(-15)^2}{35} + \frac{(-15)^2}{65}$$
$$= \frac{225}{35} + \frac{225}{65} = 6.428 + 3.46$$
$$= 9.89$$

The percentage of actual winning trades is compared with the anticipated winning trades and the losing trades with the expected losing trades. The answer must be found in the first row of Table 23.15, which gives the distribution of χ^2.

The probability is distributed unequally in the table because the results are only significant if the probability is small, showing less likelihood of the results occurring by chance. For this simple 2-element test, the result P is classified as

Highly significant   if $P \geq 10.83$ (0.1% or 1/1000)
Significant   if $P \geq 6.64$ (1% or 1/100)
Probably significant   if $P \geq 3.84$ (5% or 1/20)

TABLE 23.15   Distribution of χ^2

Cases Less	Probability of Occurring by Chance								
1	.70	.50	.30	.20	.10	.05	.02	.01	.001
1	.15	.46	1.07	1.64	2.71	3.84	5.41	6.64	10.83
2	.71	1.39	2.41	3.22	4.61	5.99	7.82	9.21	13.82
3	1.42	2.37	3.67	4.64	6.25	7.82	9.84	11.34	16.27
4	2.20	3.36	4.88	5.99	7.78	9.49	11.67	13.28	18.47
5	3.00	4.35	6.06	7.29	9.24	11.07	13.39	15.09	20.52
6	3.83	5.35	7.23	8.56	10.65	12.59	15.03	16.81	22.46
7	4.67	6.35	8.38	9.80	12.02	14.07	16.62	18.48	24.32
8	5.53	7.34	9.52	11.03	13.36	15.51	18.17	20.09	26.13
9	6.39	8.34	10.66	12.24	14.68	16.92	19.68	21.67	27.88
10	7.27	9.34	11.78	13.44	15.99	18.31	21.16	23.21	29.59

TABLE 23.16 Results from Analysis of Runs

Expected Length of Run	Actual Results (E)	Results (O)
1	1225	1214
2	612	620
3	306	311
4	153	167
5	77	67
6	38	41
7	19	16
8	10	5
9	5	3
≥ 10	4	5
≥ 8*	19	13

*The last groups were combined in order not to distort the results based on a small sample.

The answer $\chi^2 = 9.89$ is between 0.1% and 1.0%, which shows *significance*. For a large sample, the actual reliability should not have been 20% when 35% was expected.

The chi-square test can be used to compare actual price movement with random patterns to see whether there is appreciable variation. Based on the expected number of runs, discussed earlier in the section "Gambling Techniques—The Theory of Runs" in Chapter 22, Table 23.16 shows the difference between the expected runs of a random series compared to the actual runs of a price series.

Applying the actual data for runs of 1 through 8 against a random distribution, based on Table 23.16,

$$
\begin{aligned}
\chi^2 &= \sum_{i=1}^{8} \frac{(O_i - E_i)^2}{E_i} \\
&= \frac{(1214-1225)^2}{1225} + \frac{(620-612)^2}{612} + \frac{(311-306)^2}{306} + \frac{(167-153)^2}{152} \\
&\quad + \frac{(67-77)^2}{77} + \frac{(41-38)^2}{38} + \frac{(16-19)^2}{19} + \frac{(13-19)^2}{19} \\
&= 0.09877^{(1)} + 0.10457^{(2)} + 0.08169^{(3)} + 1.2810^{(4)} + 1.2987^{(5)} \\
&\quad + 0.23684^{(6)} + 0.47368^{(7)} + 1.8947^{(8)} \\
&= 5.470
\end{aligned}
$$

From Table 23.15, the probability can be found in row 7 to be about 55% for 8 cases. These results are not significant; the Theory of Runs shows that all cases taken together give the same patterns as chance movement. Individual runs or sets of 2 or 3 adjacent runs can be inspected for distortion. In both cases, the results are further from normal but not mathematically significant. The two runs that differed the most were 4 to 5 days, which showed an 11% probability of occurring by chance.

Highly significant price runs can be found in the occurrence of extended runs, for example, 20 days, which is found occasionally in trending markets. By looking at the *asymmetry* of price movement, where a reverse run of 1 day is of negligible value, the significance of these runs will dramatically increase. Price movement is not a simple matter of random runs and equal payout.

Yates Correction

Frank Yates, an English statistician, determined that the chi-square test often overestimates the significance for small amounts of data (under 5 data points). This can be corrected by subtracting 0.5 from the difference between each observed value and the expected value,

$$\chi^2_{Yates} = \sum_{i=1}^{N} \frac{\left(\left|O_i - E_i\right| - 0.5\right)^2}{E_i}$$

It should be noted that using a small number of data points is unreliable at best.

CHAPTER 24

Diversification and Portfolio Allocation

The risk of trading in specific stocks, ETFs, futures markets, or most other assets can be substantially reduced by diversification and portfolio allocation. The object of finding the right amount to invest in each asset is to allow the offsetting price movement to reduce overall risk more than it reduces returns. Trading a diverse portfolio of assets will always have less risk than a single asset, which concentrates risk. For most investors, the best portfolio will yield the highest return-to-risk ratio in real trading.

This chapter discusses the methods that can be used to build a portfolio and maximize diversification. It begins with a simple, logical selection of markets, then shows how Excel's *Solver* can provide a classic portfolio allocation solution with little effort. However, combining the return streams from trading results is not the same as using continuous returns, as is the case for a long-only portfolio of stocks. The last part of this chapter shows how a genetic algorithm approach will create a portfolio that has a better return-to-risk ratio than the traditional solutions, and is flexible in the way it handles the idiosyncrasies of active trading.

Some of the terms used in this chapter will be those used by analysts working in the field of portfolio allocation and risk management. It is important to be able to speak the same language.

Alpha, a measure of improvement above the risk-free results or the results of a benchmark. A positive alpha is a measure of being smarter.

Assets, which can be different markets, commodities, or investment vehicles.

Beta, the relative performance of a portfolio compared to the overall market or to any benchmark. A beta of 1.2 indicates that returns and volatility are 20% greater than, for example, the S&P.

Investment class, a category of investment, a similar group of markets, such as utilities, technology, equities in general, interest rates, real estate, art, or foreign exchange.

Returns, or *return series*, are the daily returns of an investment and can be net of costs, and calculated using methods shown in Chapter 2. We will normally use $E_t/E_{t-1} - 1$, where E_t is the total equity in the account as of today. Returns are effectively percent changes.

Equity is the current value of an account, including the initial investment plus or minus the cumulative net returns from trading. Where there are assets denominated in different currencies, all of them are converted to a single currency at the daily spot FX rates. We will assume U.S. dollars.

Information ratio, or *return ratio*, is the measurement used for expressing the relationship of returns to risk. Returns and risk may be calculated using different methods. This one is the annualized return divided by the annualized risk. For portfolios containing derivatives, risk may be the mean absolute semideviation, or mean absolute downside risk.[1]

Net Asset Value (NAV), an index, usually beginning at 100, that gives the value of an investment.

Objective function is the way of measuring success in a strategy. If the objective function is the return, then the method that yields the highest net profit is the best answer. If the objective function is the return ratio, then the highest return divided by risk is the best answer.

Passive investment is a buy-and-hold strategy and is used as a performance benchmark.

String, a single entry in a table, such as one possible allocation of assets. One string might assign 5% to each of 20 assets, while another string allocates 3% to each of 10 assets, and 7% to each of the remaining 10 assets.

DIVERSIFICATION

Diversification means spreading risk, the well-established way of lowering risk, and every investor is well advised to diversify. For the purposes of risk reduction, it is necessary to distinguish between *systematic risk*, which can be reduced by diversification, and *market risk*, which cannot be eliminated. The benefits of diversification are greatest when the markets traded are very different and the methods of making decisions are unrelated. Typical investment portfolios contain a variety of fixed income, equities, real estate, and art, as well as different investment philosophies. These all combine to provide different rates of return with different patterns so they do not suffer losses all at the same time. Unfortunately, this is not always the case. Market risk, including price shocks and catastrophic risk, is not predictable, and can surprise even the most sophisticated investors. The subprime crisis of 2008 proved that diversification can disappear under

---

[1] Antonio Marcos Duarte, Jr., and Maria Luiza de Andrade Maia, "Optimal Portfolios with Derivatives," *Derivatives Quarterly* (Winter 1997), 53–62. The authors prefer risk measures that are limited to downside excursions. Some risk measures were discussed in Chapter 23.

stress. Minimizing and avoiding market risk is the subject of this and other sections in this chapter.

Practical diversification can begin with a broad selection of markets and a single tactical approach, a selection of different trading strategies for one product, or a combination of multiple systems and many markets. The following sections give the primary components for achieving diversification.

Selecting Individual Assets from Unrelated Groups

Risk is reduced when trading is spread among those markets with as little relationship to one another as possible. This can be measured as a low *covariance* or low *correlation* as discussed in Chapter 6. Limiting the assets to exchange-traded stocks, futures, and funds, market groups can be roughly classified, in order of liquidity, as fixed income, currencies, equities and equity indexes, energy, industrial and precious metals, grains, livestock, foods, and miscellaneous. The last group accounts for fiber, wood, and other products. Equities are classified into many industrial groups, such as technology, health care, industrials, and utilities; however, the extent of diversification among those groups is limited because they all tend to react to the same broad economic conditions.

There is a larger category called *financials* that encompasses fixed income, currencies, and index markets. While these sectors can move independently over the short term, they are recognized as having strong fundamental interrelationships. When an economic report is released, there is a predictable pattern to the relative direction of all markets within the financials. For example, a better-than-expected announcement of U.S. GDP will first cause Treasury prices to drop (yields rise because a strong economy eventually causes inflation), then the U.S. dollar to strengthen (people buying dollars to invest in the United States), and the stock market to rise (unless it is sensitive to higher interest rates, in which case it will fall). When the attack of 9/11 occurred, investors moved away from the stock market into fixed income in a flight to safety, and away from the U.S. dollar, if possible. These patterns must be considered when allocating assets within a portfolio.

Aside from the financial markets themselves, other commodities react to changes in interest rates and the U.S. dollar. As the dollar gets weaker, the price of oil, gold, and grain increase. As actively traded world markets, these products hold an international value. If the price of crude oil is $80/bbl when 1 dollar buys 30 rubles, then a barrel of similar Russian oil will sell for 2,400 rubles in Russia. When the dollar drops to 25 rubles, down 16.7%, the price of crude oil remains at 2,400 rubles in Russia, but the U.S. price rises to $93 on the world market to maintain the same competitive price. Therefore, if interest rates directly affect exchange rates, and exchange rates directly affect the price of oil, then the price of oil will move when there is a change in interest rates or the U.S. dollar. The same is true for gold, which is quoted worldwide in U.S. dollars.

Although most news coverage comes when prices are volatile and there is correlated movement between groups, there are long intervals when prices of commodities are not driven by financials. During those periods, agricultural markets have little to do with metals, and each class of assets moves according to its own fundamentals. When the S&P

index is moving sideways, the individual stocks within the S&P will also have greater independence. Barring a price shock and its aftermath, there is considerable independent movement in most commodity prices each day, and in many stocks. However, a price shock causes all the markets to be highly correlated. Money moves the market, not fundamentals, and money seeks safety. There is no diversification during a crisis.

When selecting those markets to be used in a diversified portfolio, it is necessary to check the historical similarity of returns as well as consider the major forces that currently drive the markets. When properly selected, diversification will reduce risk, although not necessarily on any one day, as shown in Figure 24.1. As diversification increases from 1 to 4 assets, risk falls quickly; however, as the number of markets, or assets increases above 4, the reduction in risk decreases. With perfect diversification, where each investment is independent of the others, diversification into 2 assets would reduce risk to 50% of the initial risk, 3 would reduce risk to 33.3%, 4 to 25%, and so on. In the real world, a fully diversified portfolio might reduce the initial risk by 50%. More assets still offer better diversification than fewer assets.

Multiple Strategies

Given the higher degree of correlation between markets under stress, using varied strategies may give better diversification during uncertainty. Although some techniques may
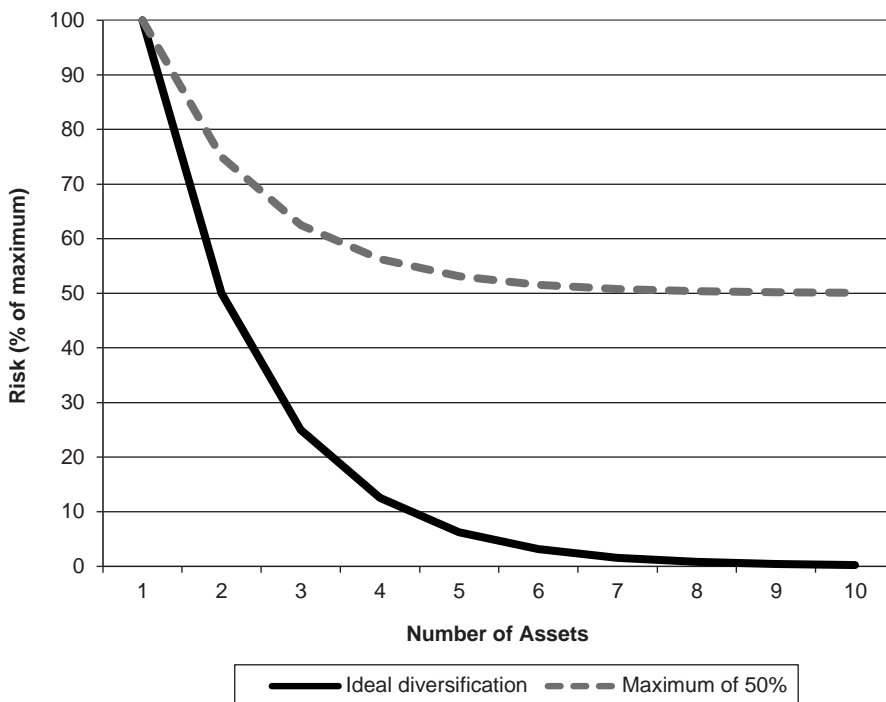
FIGURE 24.1   Effect of diversification on risk.

appear different, they may actually be highly correlated. A moving average system, ARIMA model, and point-and-figure are very different methods, but all are trend-following. If all trend-following systems are profitable, then they must extract those profits from the same price moves and hold the same position at some time. For system diversification, it is best to select strategies with different functional attributes. The following list gives methods that are likely to be less correlated, and techniques within each approach that tend to be different.

1. Trend-following (moving averages, point-and-figure, breakout).
2. Countertrend or mean-reverting (stochastic, contrary opinion, Hilbert).
3. Spreading (pairs trading, sector neutral, interdelivery, intercommodity, arbitrage, product).
4. Fundamental (value, supply and demand, P/E ratio).
5. Carry programs, convenience yield, and implied yield.

All five techniques are conceptually different. Trend-following tries to take positions in the direction of the major price move, while mean-reverting strategies target short holding periods and the opposite position. Spreads are normally neutral on direction and only benefit from the relative move between two markets, and fundamentals use value and various other statistics to take a long-term position without the benefit of timing. Carry programs and yield extraction target components of price although their methods are similar to arbitrage. These different strategies can be holding conflicting positions in the same market at any one time, reducing overall risk.

Balanced Risk

Equalizing risk is necessary if the selection of different assets and systems is to offer the best possible diversification. In securities, risk is often balanced by investing equal dollar amounts in each stock, but two stocks trading at the same price can have very different volatilities. Banking stocks that were once trading at $30 or $50 per share and have now dropped to $5 tend to be much more volatile than a stock that has always been trading at $5. Companies in the news tend to be more volatile than those that are under the radar.

The three most popular ways to allocate equal risk are

1. Investing equal dollar amounts in each asset
2. Investing equal risk as measured by the annualized standard deviation
3. Investing equal risk as measured by the average true range

If the high, low, and closing prices are available, then the average true range gives the best measure of volatility. If only the closing prices are available, then the annualized standard deviation seems to be the most popular choice. Using equal dollar amounts, which was once an industry standard, is rarely used. However, when trading stocks, you must be sure that the total value of all shares matches the amount you want to invest. For

TABLE 24.1 Finding the Number of Shares to Trade Based on Volatility

	BAC	MSFT	AAPL	Total
10/28/2011	7.05	26.98	404.95	
Ann Stdev	59.12%	23.78%	27.90%	
% of smallest	0.40	1.00	0.85	2.25
Scale	0.18	0.44	0.38	1.00
10000	1784.08	4435.45	3780.47	
Shares	253	164	9	
Equal $ shares	473	124	8	

example, if we plan to invest $10,000 in three stocks, Bank of America (BAC), Microsoft (MSFT), and Apple (AAPL), we can follow the steps in Table 24.1.

Line 1 shows the price as of October 28, 2011. Line 2 is the annualized standard deviation (volatility) of the past 200 days. Note that BAC, which has the lowest price has twice the volatility of the other stocks. Line 4 assigns a 1.0 to the stock with the lowest volatility, then finds the inverse relationship for the other stocks by dividing the MSFT volatility by the BAC and AAPL volatilities. Then a value of 0.40 for BAC indicates that it is 2.5 times as volatile as MSFT. By dividing the individual percentages on line 3 by the sum of line 3, we get the scaled percentages on line 4. Those are the percentages of the investment to be used to purchase each stock. Based on a $10,000 investment, line 5 now shows the dollar amount that can be used to invest in each stock. Dividing the number on line 5 by the price of a share (line 1), line 6 shows the number of shares that should be bought. As a comparison, the number of shares based on an equal dollar investment of $3333 are shown on line 7. Most of the risk was removed from BAC, which is highly volatile even though it is trading at a very low price. The spreadsheet with the formulas for calculating Table 24.1 are available on the Companion Website as *TSM Adjust Shares for Risk*. An additional discussion of this can be found in Chapter 23.

Diversification has its negative side as well. If assets are equally weighted, diversification gives you the average return. A classic example of diversification is the use of one strategy that has a zero net return (System *B* in Figure 24.2) but a negatively correlated
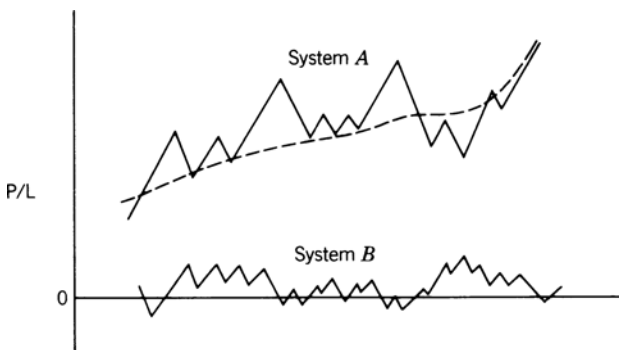
FIGURE 24.2 Improving the return ratio using negatively correlated systems.

performance to another asset (System *A*), which has a high return. While the average returns of the two assets are half the return of System *B*, the combined risk could drop by more than 50%. If the objective function is a return-to-risk ratio, then the combined portfolio of two assets is better than either one of the assets. Lower returns can be recovered by leveraging up, where possible.

CHANGING CORRELATIONS

Finding the correlation between two assets is the standard way of assessing the amount of diversification that can be achieved by allocating to both. Typically, the correlation is evaluated over a single long period of time or taken as the average of a rolling period. But averages do not explain what can happen during shorter time periods. During the fall of 2008, when the S&P plunged, all markets reversed, effectively making all correlations 1. Figure 24.3 shows the 60-day rolling correlation of the S&P against futures markets for 10-year Treasury notes, the euro, and gold. Treasury notes would normally move opposite to the S&P, that is, when the S&P rallies, interest rates would increase (T-note futures prices decrease).

On the far right, we can see that the subprime crisis caused both gold and the euro to quickly post a negative –0.6 correlation to the S&P. That is interpreted as the S&P dropping and gold and the euro rallying at essentially the same time. If we had used a shorter calculation period, the correlations would have been more extreme.

Many of the statistics that we use to make decisions are good under normal market conditions. The most common of these are correlations and value at risk. When a price shock occurs, they go to extremes for which we are always unprepared. In other parts of the book, we have discussed the only ways to avoid this unexpected risk,

- Be out of the market as much as possible.
- Cap the portfolio leverage.
- Use uncorrelated strategies.

Orderly changes in volatility can be managed using various volatility stabilization methods, discussed in the section "Volatility Stabilization."

TYPES OF PORTFOLIO MODELS

In his book, *Inside the Black Box*,[2] Rishi Narang gives four basic ways in which portfolios are constructed, along with their advantages and disadvantages.

---

[2]Rishi K. Narang, *Inside the Black Box* (Hoboken, NJ: John Wiley & Sons, 2009), Chapter 6.
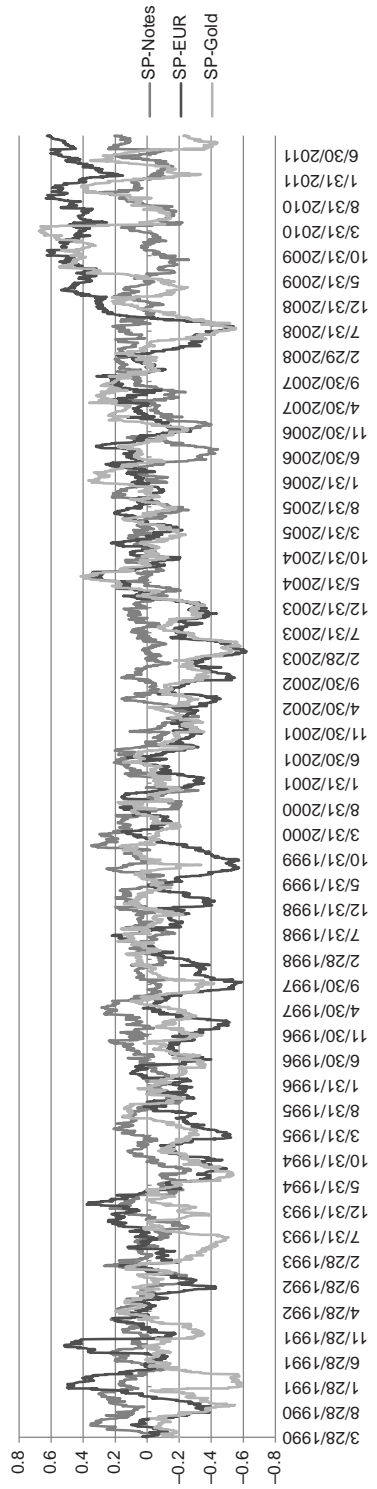
FIGURE 24.3 Rolling correlation of S&P futures against 10-year notes, the euro, and gold futures.

Equal-Dollar Weighting

Once a trade signal has been given, each asset in the portfolio receives equal weight, that is, it gets an equal share of the investment. In the case of stocks, each of 20 assets will receive 5% of the current account value. This expresses the belief that forecasting which asset will perform better than another is extremely unreliable, and that the direction of trading signal itself is the only indication of value. Unequal weighting concentrates risk, which is a problem if the trade does not meet expectations. Equal-dollar weighting is not the same as equal-risk weighting because two stocks trading at the same price can have different volatilities.

Equal-Risk Weighting

As discussed in the previous chapter, equal-risk weighting attempts to equalize the risks of each trade using a volatility measurement or estimate of drawdown. For example, if a portfolio has two assets *A* and *B*, with volatility of 4% and 1% respectively, then *A* will get 20% of the allocation and *B* will get 80%. As with equal-dollar weighting, it attempts to give every trade an equal chance to affect performance.

Alpha-Driven Weighting

In this method, those assets with higher alphas will get larger allocations; however, these are not simply scaled to the estimated alpha. Instead, the assets are typically ranked by expected alpha and those at the top given larger allocations. This avoids creating severe distortions in allocations but favors those assets that have performed best in the past.

Decision-Tree Models

A more complex approach, the decision-tree allows you to match each asset with certain characteristics to get a net allocation. For example, if there are three strategies within a portfolio, and three levels of risk under each of those strategies, then we need to associate each asset with a strategy and a risk level in order to arrive at the allocation for that asset. Either of these decisions can be dynamic.

A fifth and highly popular model is where allocations are determined by a classic *mean-variance optimization*. In this approach, the volatility and correlation of the assets are evaluated with regard to their returns, and an optimal allocation is determined. The sections that follow will discuss the various ways in which an optimal portfolio can be found.

CLASSIC PORTFOLIO ALLOCATION CALCULATIONS

Deciding which systems and markets should be traded at the same time, in an effort to reduce risk through diversification, is an entire field of expertise called *portfolio management* or *asset allocation*. While the price patterns have always evolved with structural

changes in the economy, technology, and investment participation, globalization has made portfolio management much more difficult, causing higher correlations and often greater volatility. Nevertheless, the classic approach, based on Modern Portfolio Theory (MPT), continues to be popular.

The principle of classic portfolio allocation remains sound even in a changing environment: *Diversification reduces risk.* Once the assets are selected that will be used in the portfolio, the expected return, $E(R)$, for the portfolio is the weighted average of the expected returns on the individual assets of the portfolio,[3]

$$E(R) = w_1 E(R_1) + w_2 E(R_2) + \cdots + w_n E(R_n)$$

$$= \sum_{i=1}^{n} w_i E(R_i), \quad where \sum_{i=1}^{n} w_i = 1.0$$

The worth of a portfolio is not only its return, but its risk; therefore, the covariance of the portfolio assets becomes an important criterion in deciding which combination of those assets creates the best results:

$$\sigma_R^2 = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j cov_{ij}$$

where σ_i^2 is the variance of the returns of market i, and cov_{ij} is the covariance between the returns of market i and the returns of market j. The covariance between two markets is the product of their standard deviations times the *correlation coefficient* (the *correl* function in Excel) of the same two markets

$$cov_{ij} = corr_{ij} \sigma_i \sigma_j$$

Given two portfolios with the same returns, the one with the smaller variance is more desirable; when the variances are the same, the portfolio with the highest return is best. This is the same conclusion drawn in the discussion of the *efficient frontier*. For a portfolio of only two assets, the calculations needed for expected return and variance are:

$$E(R) = w_1 E(R_1) + w_2 E(R_2)$$

$$\sigma_R^2 = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 cov_{12}$$

It is important to compare portfolios using this basic measurement of return and risk; however, it is not the entire picture. When evaluating the relationship over a long time interval, there may only be one case where a very high correlation would have caused a fatal loss of equity. That one case can be lost in the statistics because the large number of normal cases overwhelms one extreme event. For safety, it is also necessary to look at the maximum portfolio loss. In addition, it is safe to assume that the largest historic loss will be exceeded sometime in the future, if only by chance or the fact that the probability

---

[3]Peter L. Bernstein, ed., *The Portable MBA in Investment* (New York: John Wiley & Sons, 1995), 252.

of a run increases as the amount of data increases. In today's more complex market environment, the worst-case scenario becomes more likely. Globalization has increased the correlations, and the past patterns of returns are not likely to represent all of the patterns that will be seen in the near future. It is also wise to run a value-at-risk analysis on the final portfolio to see how many times risk exceeded a conservative threshold, even if it did not result in a large loss. An excessive number of risk spikes is a sign of too much leverage. Value at risk was discussed in Chapter 23.

Spreadsheet Approach

A practical way of evaluating a portfolio that does not require special mathematical knowledge or expensive software is to use a spreadsheet. Using a classic stock and bond mix, Table 24.2 shows the monthly returns of the S&P 500 (S&P) and the Lehman Brothers Treasury Index (LBTI) combined into a portfolio of 60% stocks and 40% bonds. Columns *B* and *C* show the monthly percentage returns of these two series, and columns *E* and *F* have the corresponding NAVs. At the bottom of columns *B* and *C* are some basic calculations that describe the performance. The S&P has an annualized return of 21.88% with an annualized standard deviation of 11.64%; bonds have a 9.15% annualized return with a 0.92% annualized standard deviation. The return ratio of 1.88 shows that the stock market was a better investment during these four years. Combining the two assets, which show a 0.55 correlation over these four years, gives a return of 16.46% and a ratio of 1.73, showing that the combination did not provide the diversification that was expected.

Using the standard 60% stock and 40% bond portfolio mix, shown as values in row 4, columns *B* and *C*, the combined returns are in column *D* (row 5), and can be calculated as either = B5*0.6 + C5*0.4 or = sumproduct(B$4:C$4,B5:C5). Columns *E*, *F*, and *G* are the NAVs, calculated in the standard way, where cell E7 is = E6*(1 + B7/100) changing the returns from whole numbers to percentages. During this period, the stock market was moving steadily higher ahead of what was to be the Internet bubble. If the data used was 10 years later, bonds would have outperformed stocks. This simple method can be used for any number of assets, provided you know the allocations that are assigned. The spreadsheet for Table 24.2 can be found on the Companion Website as *TSM Stock and Bonds Portfolio*.

FINDING OPTIMAL PORTFOLIO ALLOCATION USING EXCEL'S SOLVER

Excel's *Solver* is a remarkably simple tool for solving complex problems.[4] If it cannot be found under *Tools*, then it will be necessary to select *Add-Ins* and follow the instructions for installing this software. It is part of the Excel package, but it is not set up during the initial installation.

---

[4]The author would like to thank Evans Curtice, of Graham Capital Management, for his help in this section.

TABLE 24.2 Portfolio Evaluation of Stocks and Bonds Using a Spreadsheet

Month	Percentage Change			NAVs		
	S&P	LBTI	Stock + Bond	S&P	LBTI	Stock + Bond
A	B	C	D	E	F	G
Allocation	**0.60**	**0.40**		100.00	100.00	100.00
Jan-94	3.40	2.44	3.02	103.40	102.44	103.02
Feb-94	−2.71	−4.03	−3.24	100.60	98.31	99.68
Mar-94	−4.00	−4.22	−4.09	96.57	94.16	95.61
Apr-94	1.28	−1.36	0.22	97.81	92.88	95.82
May-94	1.64	−0.61	0.74	99.41	92.32	96.53
Jun-94	−2.45	−0.09	−1.51	96.98	92.23	95.07
Jul-94	3.60	3.03	3.37	100.47	95.03	98.28
Aug-94	4.10	−0.53	2.25	104.59	94.52	100.49
Sep-94	−2.45	−3.09	−2.71	102.03	91.60	97.77
Oct-94	2.25	−0.43	1.18	104.32	91.21	98.92
Nov-94	−3.64	0.48	−1.99	100.52	91.65	96.95
Dec-94	1.48	2.42	1.86	102.01	93.86	98.75
Jan-95	2.59	2.68	2.63	104.65	96.38	101.34
Feb-95	3.90	2.88	3.49	108.74	99.16	104.88
Mar-95	2.95	0.64	2.03	111.94	99.79	107.01
Apr-95	2.94	1.78	2.48	115.24	101.57	109.66
May-95	4.00	7.90	5.56	119.84	109.59	115.76
Jun-95	2.32	1.06	1.82	122.62	110.75	117.86
Jul-95	3.32	−1.64	1.34	126.70	108.94	119.43
Aug-95	0.25	2.13	1.00	127.01	111.26	120.63
Sep-95	4.22	1.73	3.22	132.37	113.18	124.52
Oct-95	−0.36	2.93	0.96	131.90	116.50	125.71
Nov-95	4.39	2.57	3.66	137.69	119.49	130.31
Dec-95	1.93	2.67	2.23	140.34	122.68	133.21
Jan-96	3.26	−0.02	1.95	144.92	122.66	135.81
Feb-96	0.69	−4.82	−1.51	145.92	116.74	133.75
Mar-96	0.79	−1.85	−0.27	147.07	114.59	133.40
Apr-96	1.34	−1.72	0.12	149.04	112.61	133.55
May-96	2.28	−0.04	1.35	152.44	112.57	135.36
Jun-96	0.22	2.10	0.97	152.78	114.93	136.67
Jul-96	−4.57	0.00	−2.74	145.79	114.93	132.92
Aug-96	2.11	−1.34	0.73	148.87	113.39	133.89
Sep-96	5.63	2.74	4.47	157.25	116.50	139.88
Oct-96	2.76	3.94	3.23	161.59	121.09	144.41
Nov-96	7.56	3.32	5.86	173.81	125.11	152.87
Dec-96	−1.98	−2.34	−2.12	170.37	122.18	149.63
Jan-97	6.25	−0.84	3.41	181.01	121.16	154.73
Feb-97	0.78	−0.01	0.46	182.43	121.14	155.45
Mar-97	−4.10	−2.44	−3.44	174.95	118.19	150.11
Apr-97	5.97	2.34	4.52	185.39	120.95	156.89
May-97	6.09	1.14	4.11	196.68	122.33	163.34
Jun-97	4.48	1.94	3.46	205.49	124.71	169.00
Jul-97	7.96	5.89	7.13	221.85	132.05	181.05
Aug-97	−5.60	−2.87	−4.51	209.43	128.26	172.89
Sep-97	5.48	2.84	4.42	220.90	131.90	180.54
Oct-97	−3.34	3.34	−0.67	213.53	136.31	179.33
AROR	21.88%	8.42%	16.46%			
ASTD	11.64%	9.15%	9.50%			
Ratio	1.88	0.92	1.73			

Solver allows the user to solve equations, subject to various conditions. For this example, we would like to know how we should have allocated our investment during the past year to a portfolio of S&P 500 futures, gold futures, Apple, and Microsoft. The steps are:

1. Align the data as in columns *B–E*, Table 24.3, starting in row 4. When using mixed assets, such as NYSE stocks and gold futures, there may be days when the NYSE was open, but gold closed. On those days, the previous price of gold is carried forward. Calculate the returns of columns *B–E* in columns *F–I* using $p_t/p_{t-1} - 1$. Returns start in row 5. Figure 24.4 shows the original prices used. MSFT is shown alone on the right scale because its value is much lower than the other three members of the portfolio.

2. Enter the initial, equally-weighted allocations in row 3, columns *F–I*, above the returns.

3. Enter the formula for the sum of those allocations (always 1.0) in row 4, column I.

4. Calculate the weighted returns of the four assets in column *J*.

5. Calculate the NAVs in column *K*.

6. At the top of column *J* calculate the annualized standard deviation of the returns in column *J*.

7. At the top of column *K* calculate the annualized rate of return from column *K*.

8. In cell *L*3 calculate the information ratio cell *J*3/*K*3. This is the value that will be used to determine the success of the solution.

The value of the returns and volatility, 20.0% and 14.6%, give a good ratio of 1.36, shown in row 3, columns *H–J*. The Solver page can now be set up.

Solver Settings

By clicking on *Data/Solver* (Excel 2010) or *Tools/Solver* (earlier versions), the Solver window will open. This has the following sections that must be filled in:

1. *Set Objective* (or *Target cell*), a single value that will be used to decide the solution. In this example, the target cell is *L*3, the information ratio.

2. *To* (or *Equal to*) allows you to say that the maximum value (max) of the information ratio is needed.

3. *By changing cells*, tells Solver that the solution can be found by varying the values of the weighting factors in $F$3:$I$3.

4. *Subject to the constraints*, allows us to enter the requirement that the sum of the weighting factors must equal 1.0. This is done using a secondary window where the cells, mathematical operator (=, <, >), and a constant can be entered. This portfolio requires the cell $I$4, the operator "=", and the constant 1. It also requires that each

TABLE 24.3 Solver Set-Up

A	B	C	D	E	F	G	H	I	J	K	L
		Prices					Returns			Portfolio	
Date	S&P	GOLD	AAPL	MSFT	S&P	GOLD	AAPL	MSFT	Returns	NAVs	Ratio
Weights					0.250	0.250	0.250	0.250	14.6%	20.0%	1.3651
10/29/2010	1158.75	1366.3	300.87	25.84				1.000		100	
11/1/2010	1162	1359.3	302.2	26.06	0.0028	−0.0051	0.0044	0.0085	0.0027	100.27	
11/2/2010	1171.75	1365.6	307	26.38	0.0084	0.0046	0.0159	0.0123	0.0103	101.30	
11/3/2010	1176.25	1346.3	308.53	26.32	0.0038	−0.0141	0.0050	−0.0023	−0.0019	101.11	
11/4/2010	1197.5	1391.8	315.03	26.37	0.0181	0.0338	0.0211	0.0019	0.0187	103.00	
11/5/2010	1201	1406.4	316.75	25.89	0.0029	0.0105	0.0055	−0.0182	0.0002	103.01	
11/8/2010	1199	1411.9	316.76	25.94	−0.0017	0.0039	0.0000	0.0019	0.0011	103.12	
11/9/2010	1190	1418.8	314.5	26.07	−0.0075	0.0049	−0.0071	0.0050	−0.0012	103.00	
11/10/2010	1193	1408	313.55	26.17	0.0025	−0.0076	−0.0030	0.0038	−0.0011	102.89	
11/11/2010	1190	1412	314.25	25.64	−0.0025	0.0028	0.0022	−0.0203	−0.0044	102.44	
11/12/2010	1174.5	1374.2	303.63	25.46	−0.0130	−0.0268	−0.0338	−0.0070	−0.0202	100.37	
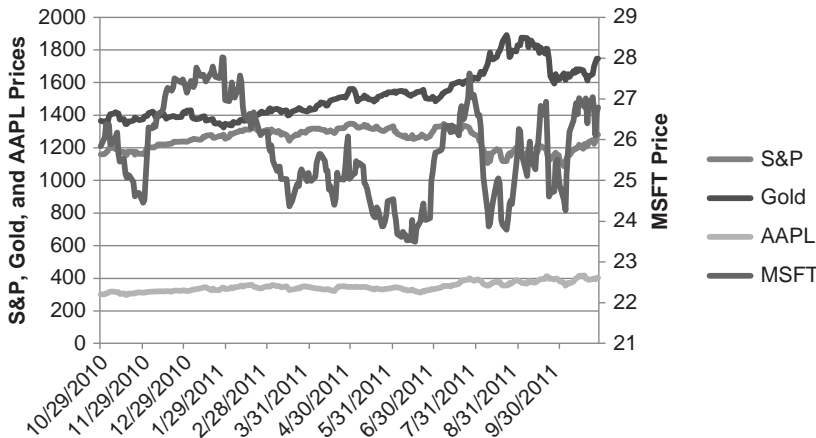
FIGURE 24.4   Prices of the four assets used in the Solver solution.

of the cells F3, G3, H3, and I3 be set so that each cell is "> = 0". This will prevent negative weighting factors, which represent selling the asset rather than buying. If the *short sale solution* was desired, then each of the 3 cells would be constrained as, for example, F3 > = −1 and F3 < = 1. This would keep the weighting factors between −1 and +1; otherwise they can take on any value.

There are various options for the number of iterations to be performed, the technique used, the maximum time allowed, and a record of various calculated values. The defaults work well; therefore, none of these are necessary.

Solver Results

By clicking on the *Solve* button, the program returns an instantaneous answer. The cells containing the allocations are changed, as are the portfolio returns, AROR, and information ratio, as shown below. The ratio has improved from 1.3651 to 1.9723 by heavily weighting gold and Apple and allocating nothing to Microsoft. A plot of the original and optimal NAVs are shown in Figure 24.5. Of course, this is all based on knowing what happened during the past year. The much more difficult problem is deciding if these allocations are likely to perform well in the following year.

Returns (New Allocations)				Portfolio		
S&P	**GOLD**	**AAPL**	**MSFT**	**Returns**	**Risk**	**Ratio**
0.091	0.550	0.359	0.000	15.2%	30.0%	1.9723

Using more data and a broader set of assets will always give a more robust solution, but the solution assumes that these same allocations, that is, the same markets, will continue to perform in the same way, relative to one another, as they did in the past. That
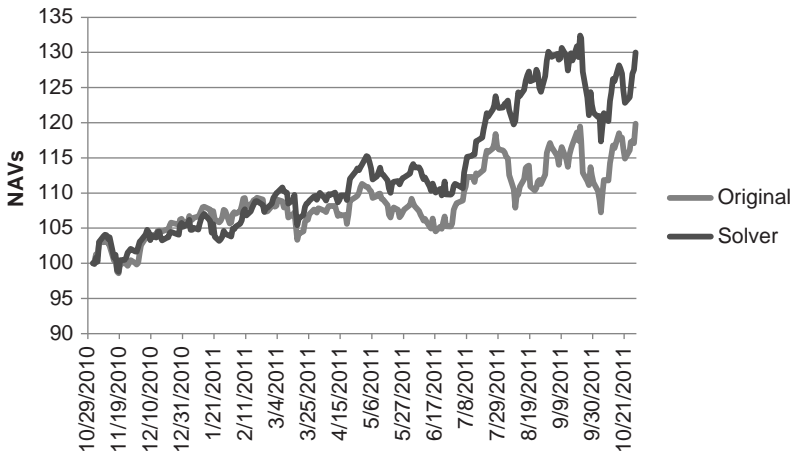
FIGURE 24.5   Results of optimal portfolio compared to the original equally weighted portfolio.

is not clear. In an attempt to stay in tune with market performance, many analysts will recalculate the allocations periodically, using a limited amount of past data. One may use the past 5 years and recalculate quarterly, while someone else will use 3 years and recalculate monthly. There are no studies that show which of these, if any, has the best forecasting ability.

KAUFMAN'S GENETIC ALGORITHM SOLUTION TO PORTFOLIO ALLOCATION (GASP)

A portfolio allocation is a basic requirement for trading more than one market. Excel's Solver can be used to get an estimate of the proper allocations; however, it does not include the cross-correlations in price movement, market liquidity, or a number of subtle measurements that are needed by the professional manager handling large portfolios.

Portfolio optimization (specifically, *mean-variance optimization*) is the accepted method used by many financial analysts to find the best allocations. It has been taught in university graduate school financial courses as the recommended solution. Many Wall Street firms apply this method and *rebalance* the portfolio quarterly, using as little as one year of historic prices. This gives the allocation that maximizes the return ratio, and places their expected returns on the optimal point on the efficient frontier.

Whether the need is to find a portfolio of stocks that work well together, balance a diverse mix of futures markets, or balance any combination of assets, it will require a special computer program. Classic portfolio optimization deals with choosing the set of stocks that fit best together, and positions are always long. It is called a *passive portfolio* because there is no strategy involved, simply an allocation. Once a position is entered, it

is held for a long time, until the next optimization is run or the portfolio is rebalanced. All of the investment is fully used.

This section discusses the major considerations and their solutions when the assets are trading results of different stock and futures markets, rather than stock prices or other cash data series. With trading results, there are periods when a strategy has no position; therefore, there are no returns. As we will see, this type of input presents a problem for the traditional mean-variance method, such as used by Markowitz, which assumes that there will always be prices, hence returns, each day. However, it is easily solved using the *Genetic Algorithm Solution to Portfolios* (GASP) developed by Kaufman.

Modern Portfolio Theory

Chapter 20 gave a basic explanation of how to use a genetic algorithm to find the parameter values of a trading system. Here, the genetic algorithm is used to find the best allocation for each asset—stock, stock index, or futures market—in order to achieve the best objective function.

The problem of portfolio allocation is different for the active trader than for the passive investor. In Modern Portfolio Theory,[5] Markowitz shows how optimal diversification can be achieved by allocating a percentage of an investment to each stock using a mean-variance solution. Because of Markowitz and those who followed him, it is generally accepted that a properly allocated, risk-adjusted portfolio of assets is needed to achieve the highest return for a given level of risk. However, Markowitz did not address the problem of using trading returns as a substitute for continuous prices. Returns from trading present unique problems.

A Special Case for Active Trading

For those traders who use a systematic method of buying and selling, their returns reflect those decisions. For them, an optimum portfolio based on the prices of the markets traded makes no sense. The portfolio program may be treating two stocks that are rising at the same time as highly correlated, while the trading system shows one of those stocks long and the other short. Instead of a strong positive correlation for that time period, there should be a strong negative correlation. Therefore, it is the daily returns of the trading in each market that reflect the positions, not the price. Classic portfolio optimization uses returns that are always calculated as the change in price over the observation period (for example, daily or monthly).

---

[5] Harry M. Markowitz, "Portfolio Selection," *Journal of Finance* (1979), also Harry M. Markowitz, *Portfolio Selection* (New Haven: Yale University Press, 1959). The Markowitz model allows for short sales represented by negative weighting factors, where the sum of all weighting factors is 1. In this genetic algorithm, there will be only positive allocations, although negative and positive can be implemented by allowing a random number between –1 and +1 to represent the weights rather than a value between 0 and 1.

In traditional portfolio analysis, called a *mean-variance solution*, the standard deviation of the daily returns is used to represent risk. The covariance of each market with respect to another, taken as pairs, represents the correlation and relative volatility of one with respect to the other. This method is accurate for continuous price series, where the daily price changes represent returns. However, when using trading returns instead of continuous prices, the method fails.

For example, when trading the S&P and bonds, a systematic method that uses two moving averages has the following rules that create longs, short sales, and neutral situations:

1. Hold longs when the fast average and the slow average are both moving higher.

2. Hold shorts when the fast average and the slow average are both moving lower.

3. Exit all positions and do not enter new ones when the fast and slow averages are moving in different directions.

Problems with the Standard Deviation Calculation

The previous trading rules result in a system that may be in the market only 50% of the time. During the periods when there are no positions, the daily returns are zero. How does this affect the standard deviation? Let us assume that a bond price series (as well as a passive holding in bonds) has a daily standard deviation of 0.5%. If the trading system just described is in the market for 50% of the time, then the standard deviation will be only 0.25% because one-half the days have a zero return

Because the standard deviation is used to represent risk, a trading method that is in the market less often will always appear to have less risk, and we would allocate larger positions to it. For example, if the average trading returns from stock A and stock B were both the same but stock A was in the market 50% of the time and stock B was in 25% of the time, then the standard deviation for stock B returns would be 50% of the stock A returns. It would therefore get positions that were twice the size of stock A. But, on a day when they are both in the market, they both have the same risk. By doubling the size of stock B we have distorted the portfolio rather than balanced it, and have more risk than we expected. Using the classic method of allocation, we would trade larger positions in the market that is least active, that is, the one with the most zero returns. To be fair, some analysts would argue with this conclusion and say that the zero returns count, and any two series with the same standard deviation have the same risk. Perhaps they should use their allocations to trade their own portfolio.

The Problem with the Covariance Calculation

The same intermittent returns are a problem for the covariance calculation. Embedded in this method are the covariances between every pair of assets. If the data are prices and they move up and down together, then the covariance between the two will be small; if they move differently, the covariance will be greater. Suppose that the S&P and bonds strategies were each in the market 50% of the total days, and that they were in at the same

time only 25%. How do you measure the similarity of returns when, in 75% of the days, there were no returns for one of the two markets?

Consider a more extreme case, where two strategies, one trending and one mean-reverting, are applied to the same market. When the market is trending, the trend system is used, and when the market is sideways the mean-reverting system is used. The two strategies both hold positions on only 3% of the total days, yet a correlation coefficient, or a covariance, is calculated using the returns of every day. The combined returns when both strategies are trading would be overwhelmed by the far larger set of returns when only one is trading. This would cause a jump in risk during actual trading that may not have been accounted for. Yet it would be statistically correct.

The standard deviation and covariance seem to miss the intent of the measurement, which is to represent the risk of the two assets when they are active at the same time. Modern Portfolio Theory arrives at a definitive, *optimal* allocation for any portfolio of assets. If, by applying a different technique using a genetic algorithm, the solution gives a return ratio higher than the mean-variance solution, we can show empirically that there must have been some problem in the calculations when applied to intermittent trading returns.

Other Problems with Infrequent Trading Activity

Some trading methods may be very selective, waiting for the right combination of patterns to enter a position. When trading activity is low, for example, positions are held only 10% to 20% of the time, the cumulative equity series appears to have alternating periods of sharp increases (when generating trading profits) and a horizontal line (when out of the market). The net result is that the annualized returns are quite low relative to the total portfolio, and the returns have a high standard deviation relative to the rate of return. However, these strategies have a high return when they are in the market and offer important diversification to the portfolio. When applying a traditional mean-variance portfolio solution, these strategies and returns are eliminated because they have low betas (returns relative to an index of all systems), even though they have the highest payout (returns per day) when they are in the market during highly volatile periods.

The Objective Function and Target Risk Level

Before the problem can be solved, it is necessary to decide how to measure success. The accepted approach is the return ratio, or information ratio, the annualized returns divided by the annualized risk; however, there are many ways to measure risk.

> *Risk measured by annualized volatility.* Securities analysts have always measured risk as the annualized volatility, the standard deviation of daily, monthly, or annual returns (if there is enough data), multiplied by the square root of the number of periods in a year, 252 for daily, 12 for monthly. Using daily returns, which is most common, requires that we accept the transformation to annualization using $\sqrt{252}$. Although this is a long-standing, accepted method, there should be some concern that estimating annualized returns from daily returns can be inaccurate.

Risk measured by the residuals of a linear regression. Another popular measure of risk uses the residuals of a linear regression. This can be derived from the same calculation that gets the annualized rate of return of the portfolio based on the slope of the straight-line approximation. If there are 5 years of return history, then the residuals would give the standard deviation of the return differences above and below the straight-line fit for the 5 years. This results in a probability of loss relative to the 5-year history, a convenient and useful value, which does not need to be annualized as did the standard deviation.

    The negative aspect of using residuals as a risk measurement is that the optimization process will find the best allocation that keeps returns as close to the straight line as possible. This may seem good when it prevents drawdowns, but it also penalizes any prolonged gains that move returns far above the normal returns, as represented by a straight line.

Semivariance of daily drawdowns. One measure that stands out as not having the restrictive qualities of either the standard deviation or the residuals is the semivariance of drawdowns. This measure uses the difference between the previous high cumulative return and the current cumulative return, always as a positive number. Therefore, there is a non-zero value on every day that is not a new high equity value. There is no accounting for the pattern of profits, only the losses. Because this represents the probability of a drawdown of a specific size over the full interval of returns, it is not necessary to annualize the results. The portfolio optimization process will try to limit the size of the drawdowns, but has no agenda with respect to the profit patterns—they do not have to be close to a straight line, nor do they need to be small.

The objective function (OF) used in GASP, the genetic algorithm solution, will then be

$$OF = AROR/\sigma_D$$

where σ_D = the standard deviation of the daily drawdowns, or semivariance,

$$\sigma_D = \sigma(E_H - E_i),\ i = 1,\ n$$

where E_H = the high equity as of day i
          E_i = the current equity
          $E_H \neq E_i$
          n = the total days of equity

The total items used for the standard deviation will be less than n by the number of days on which there was a new high equity, $E_H = E_i$.

    The *target risk level* is the probability of a specific loss, given as a standard deviation or a percentage of investment, over a specific period of time. You may specify that the portfolio must have a small chance, for example, 0.1% (one-tenth of 1%) of losing 10% during a 10-year period. The 10% loss is based on the measurement of risk, either the standard deviation of the returns, possibly the standard deviation of drawdowns

(semivariance), or the Ulcer Index (see Chapter 23 for different risk measures). Both the maximum and minimum target risk level will need to be specified. The minimum is necessary because the program can always allocate a very small part of the investment in order to keep risk low. That will normally cause the rate of return to be unacceptably low. When searching for the optimal portfolio, a large number of unacceptable solutions can be avoided by creating a range of risk tolerance.

The Optimal Portfolio

The *optimal portfolio* is the same whether using the classic mean-variance solution or a genetic algorithm approach. It is always the allocation with the best objective function, in our case, the highest return for the lowest possible risk, also called the *information ratio*.

Without considering investor risk preference, where some investors will accept higher risk in exchange for a smaller marginal return, any rational investor will take:

- Higher returns for the same risk.
- Lower risk for the same returns.

The efficient frontier was introduced in Chapter 22, showing the curve that represented the best choices of portfolios that satisfied the objectives of the rational investor. Figure 24.6 shows the efficient frontier with the optimal portfolio marked as point *A*. If
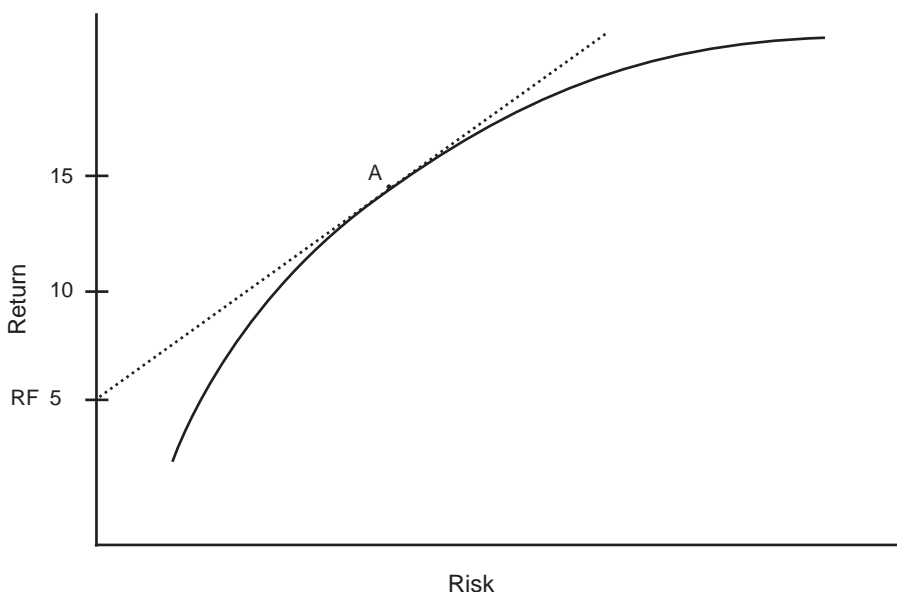
FIGURE 24.6   The optimal portfolio located on the efficient frontier.

the risk-free rate of return (RF) is 5.0%, then the optimal portfolio is located at the point where a straight line, drawn from point *RF*, touches the efficient frontier. This will also be the point where the return ratio is largest. Portfolios on the efficient frontier that are higher and to the right of *A* have risk increasing faster than returns; those that are lower and to the left of *A* have returns declining faster than risk. Point *A* will move farther to the right as the risk-free rate of return increases.

Portfolio Constraints

Portfolio constraints represent practical considerations and limitations that are imposed on the portfolio program. The most obvious of these constraints is liquidity. Some stocks or futures markets are more active than others. The portfolio optimizer cannot be allowed to allocate too much to a thinly traded stock. That would result in slippage and a poor execution far in excess of those transaction costs considered in the testing of the strategy. A safe rule is not to exceed 1% of the average volume traded in one day; some managers would say 3%. Trading less is always better.

It is practical to specify liquidity constraints as absolute numbers, such as 500,000 shares of Microsoft or 5,000 contracts of S&P minis. Restricting allocation to a percentage of the investment may not always translate into the right number of shares or contracts.

It may also be necessary to *restrict negative allocations*, or weights, which imply that you would take the opposite positions in those assets. For example, a portfolio that has both euros and gold may find that euros have a better return ratio than gold and appreciate faster. The ideal allocation may be to place 5% of the investment into euros and –3% into gold, effectively arbitraging euros and gold to net a better ratio. Some investors may be fine with this solution; however, most investors would prefer only positive allocations.

Constraining Diversification

Ideally, the portfolio allocation program should be allowed to mix and match any markets in any percentage to create the best portfolio. However, in a real world there are choices to be made. There is a practical reason for choosing between

1. Optimizing a portfolio of all stocks at one time.
2. Optimizing the returns of stocks within a sector or industrial group, then optimizing the returns of the groups using the allocations from the first step.

Although a genetic algorithm solution is fast, time is always a factor. Separating the assets into common groups, optimizing these smaller sets and creating a subportfolio return series, then optimizing the results of those smaller portfolios turns out to be a much faster process than a single, large optimization. Because the time to calculate expands exponentially, 10 separate optimizations each with 10 assets is much faster than

one optimization with 100 assets. In addition, we have the opportunity to impose some degree of common sense on a purely mechanical process.

We may know more about the potential for correlated price movement than the portfolio program can measure using the return streams, having lived through the past few years of turmoil. All assets have the ability to be highly correlated in times of extreme stress. If the price history does not have enough occurrences of the type of events that cause common price movement, then the portfolio process will assign allocations to similar assets that may be too large. When finding a portfolio allocation using only a recent time interval, the correlations may not appear strong; however, infrequent events that move these stocks in the same direction are very important to risk. There should be enough data, and it should be selected to include the high-risk events.

The Genetic Algorithm Solution

The GASP program uses a genetic algorithm solution because it is fast, provides a non-linear approach, and can handle disjoint returns, common to trading performance. There are six steps in the genetic algorithm solution:

1. Initialization
2. Calculation of the portfolio objective function
3. Test for completion
4. Propagation
5. Mating
6. Mutating

After Step 6, the process returns to Step 2 and ends when Step 3 is satisfied.

Preparation

To begin the process of portfolio allocation, we first need the return series of each asset. In a traditional stock portfolio, these can be the daily returns derived from the price series using either $p_t/p_{t-1} - 1$, or $\ln(p_t/p_{t-1})$. For traders, in particular futures traders, the return series are the daily profits or losses of each strategy applied to each market, converted to a percentage change based on a designated investment which reflects the accumulated profits and losses. Data may also be input as NAV series, which are then easily converted to daily returns. Table 24.4 shows 5 assets representing the NAVs of one strategy applied to 5 different markets, for example AAPL, GE, the euro futures, 10-year Treasury note futures, and crude oil futures. This data history begins arbitrarily on 8/23/2000.

Create common trading units. The final portfolio will assign a percentage of the investment to each asset based on the return series. That percentage should be applied to the assets in the same way that the original returns were created. This is straightforward

TABLE 24.4   Returns Expressed as NAVs

Step 1. Convert Input Returns to NAVs

Date	A	B	C	D	E
8/23/2000	100.00	100.00	100.00	100.00	100.00
8/24/2000	100.00	99.24	99.09	98.83	99.15
8/25/2000	99.93	100.02	100.02	100.03	100.02
8/28/2000	101.66	101.66	101.98	102.55	101.86
8/29/2000	101.39	101.40	101.33	102.16	101.58
8/30/2000	100.35	102.54	101.25	103.90	102.85
8/31/2000	99.94	102.15	101.06	103.30	103.33
9/1/2000	100.47	101.58	101.06	102.43	103.33
9/5/2000	101.40	102.00	101.56	102.20	103.80
9/6/2000	102.45	104.09	103.82	105.41	105.26
9/7/2000	102.45	104.67	103.82	106.86	105.26
9/8/2000	102.45	104.67	103.82	101.40	105.26
9/11/2000	101.22	104.05	103.88	100.52	104.72
9/12/2000	99.45	102.38	103.88	97.94	102.84
9/13/2000	98.50	101.48	103.88	96.56	101.83
9/14/2000	99.01	100.66	103.88	95.60	100.91
9/15/2000	102.83	101.45	105.03	96.30	100.91
9/18/2000	104.48	103.07	106.63	99.30	100.91

for a portfolio of only stocks, but when mixing stocks and futures, you will need to know the initial investment allocated to futures to assign the new allocation correctly.

If the returns of the assets were generated with different investments, the results will try to equalize the risk. For example, if the return series for asset A was based on a \$2 million investment, and the other four assets, B through E, were each based on \$1 million, then a final portfolio allocation would assign 11.1% to A and 22.2% to each of the other assets in order to create equal investment risk. It recognizes that the volatility of the returns of the \$2 million asset is twice that of the other assets. Asset A will then be traded with half the amount as the other assets, which will make the net amounts equal for all assets.

To avoid confusion, it is easiest to assign equal, even-value investment amounts when creating the returns for each system or market.

Initializing the Pool

The search for the best portfolio allocation begins by creating a *pool* of random portfolios, commonly called *strings* in genetic algorithm terminology. The pool can also be thought of as a *gene pool*, because future portfolios will be created from it. The pool should be as large as possible in order to increase the speed of the search. For portfolios with fewer than 25 assets, a pool of 5,000 strings should be large enough. To understand the number of combinations in a complete, sequential search of all portfolios, consider

the 5-asset example, which has 100^5, or 10 billion, combinations—and that is if the percentage allocations are in whole percent, not fractions of a percent.

Creating a Portfolio from Random Numbers

Using the 5-asset example, one portfolio, or string, is created by:

1. Finding a *uniform* random number between 0 and 1 for each of the 5 assets. In Table 24.5, the Excel function RAND( ) is used for each of the assets *A–E*, and for each of the portfolios, 1–18 in the pool.

2. Adding the 5 random numbers for each portfolio to get the total, *T*, shown in the right column.

3. Dividing each of the random numbers for assets *A–E* by *T* to create a normalized percentage, where the total of all random numbers equals 1. These numbers represent the weighting factors, or initial allocations for each asset, shown in Table 24.6.

4. *Applying constraints.* Not all portfolios will qualify. If the investment is large—$100 million, for example—and asset *A* is a stock that must be limited to purchases less than $10 million, then asset *A* cannot have a weighting factor greater than 10%. This can be resolved by rejecting any portfolio that shows a weighting factor greater than 0.10 under asset *A*. In Table 24.6 that would reject all but two portfolios, 7 and 16. Fortunately, the process of generating initial portfolios is very fast, and

TABLE 24.5   Generating Random Numbers to Create Weighting Factors

Step 2. Generate Random Numbers between 0 and 1

Portfolio	A	B	C	D	E	Total A–E
1	0.954	0.356	0.402	0.924	0.388	3.025
2	0.642	0.091	0.501	0.010	0.010	1.255
3	0.295	0.995	0.509	0.558	0.687	3.044
4	0.581	0.204	0.943	0.408	0.175	2.312
5	0.615	0.796	0.277	0.115	0.072	1.874
6	0.726	0.073	0.069	0.376	0.184	1.428
7	0.953	0.523	0.246	0.893	0.080	2.695
8	0.329	0.667	0.169	0.211	0.626	2.001
9	0.660	0.747	0.372	0.226	0.769	2.774
10	0.833	0.893	0.536	0.010	0.850	3.123
11	0.776	0.252	0.663	0.964	0.667	3.321
12	0.578	0.873	0.299	0.100	0.447	2.296
13	0.883	0.567	0.621	0.712	0.732	3.515
14	0.606	0.025	0.798	0.227	0.219	1.875
15	0.629	0.396	0.234	0.041	0.146	1.446
16	0.713	0.531	0.515	0.852	0.942	3.553
17	0.829	0.584	0.411	0.289	0.011	2.125
18	0.812	0.215	0.444	0.159	0.724	2.354

TABLE 24.6   Normalized Weighting Factors for Each Portfolio in the Pool

Step 3. Normalize Random Values

Portfolio	A	B	C	D	E	Total A–E
1	0.131	0.290	0.062	0.310	0.208	1.0
2	0.150	0.125	0.291	0.290	0.145	1.0
3	0.393	0.103	0.295	0.135	0.074	1.0
4	0.304	0.130	0.132	0.385	0.049	1.0
5	0.248	0.287	0.058	0.160	0.247	1.0
6	0.330	0.171	0.280	0.057	0.163	1.0
7	0.066	0.313	0.201	0.203	0.217	1.0
8	0.380	0.043	0.143	0.423	0.012	1.0
9	0.216	0.365	0.133	0.016	0.270	1.0
10	0.277	0.049	0.129	0.261	0.284	1.0
11	0.178	0.283	0.301	0.107	0.130	1.0
12	0.172	0.086	0.204	0.288	0.250	1.0
13	0.439	0.322	0.053	0.172	0.014	1.0
14	0.305	0.063	0.063	0.217	0.352	1.0
15	0.313	0.189	0.245	0.251	0.002	1.0
16	0.098	0.247	0.359	0.255	0.041	1.0
17	0.111	0.322	0.079	0.251	0.237	1.0
18	0.214	0.165	0.225	0.139	0.257	1.0

rejecting 90% of them will not slow the process significantly. It would seem more efficient to generate a random number for asset *A* using the Excel function *Rand-between*(0,0.10) to produce a value between 0 and .10; however, the rescaling after the other random numbers have been generated might result in a value for *A* greater than 10%.

Group liquidity constraints must also be corrected. After each asset has been assigned an allocation, another random number is selected between zero and the maximum liquidity for an entire group. Then the assets within that group are scaled to that value. Consider a group with liquidity set at 1,000 futures contracts. This group contains three assets that have been assigned maximum liquidity of 550, 100, and 300 contracts based on their individual system liquidity constraints; however, the total allocation for the group that contains this system may be less than the group maximum. A random number is then selected for the group liquidity. If, for example, the number 0.40 is selected and the maximum group liquidity is 1,000, then all allocations in that group will be rescaled to total 400. When rescaling, it is necessary to prevent the individual assets from exceeding their maximum liquidity.

Both *maximum and minimum absolute risk* is another constraint. Although the objective function will be returns divided by risk, the amount of absolute risk will be important in the results. Some investors will not accept a portfolio in which the risk

exceeds some predetermined value. In other cases, it is not possible to achieve returns greater than, for example, 20% per annum, without having annualized risk in the range 8% to 20%. Any random portfolio that produces a risk less than 8% or greater than 20% will be rejected. The calculations for finding these values are in the next section.

Evaluating Each String in the Pool

The evaluation process calculates the portfolio returns of each string and analyzes the results in order to find the objective function. This is done immediately after the initial pool is created, and it is done at the beginning of each cycle throughout the genetic algorithm process. The objective function chosen for this test is the information ratio; therefore, it will only be necessary to find the annualized returns and annualized standard deviation of each portfolio in the pool.

The portfolio returns R on any day, i, are

$$R_i = w_1 R_{1i} + w_2 R_{2i} + w_3 R_{3i} + w_4 R_{4i} + w_5 R_{5i}$$

where $w_1 + w_2 + w_3 + w_4 + w_5 = \Sigma w = 1$, and $0 \leq w_i \leq 1$

R_i = the portfolio return on day i

r_{ni} = the daily return of asset n on day i

This process is easily done on a spreadsheet, shown in Table 24.7 using either returns or, in this case, the NAVs. For the first row, on 8/23/2000, each of the NAVs is multiplied by the weighting factor at the top of the column (using *sumproduct*) and is posted to that row to the right. The final column shows the returns of that day. This is done for each day to get a portfolio NAV series. Once the daily returns are created for the first portfolio (or any of the 5,000 strings in the pool), the annualized returns and annualized standard deviation are calculated, shown as 1.85% and 1.23% per year. At the bottom of the column *Portfolio NAV* is the AROR of 29.33% for these 18 days. At the bottom of the Returns column is the annualized volatility of 19.15%, giving an information ratio (the objective function OF_1) of 1.53 for Portfolio 1.

To complete the evaluation step, the objective functions, OF_p, must be calculated for all p randomly generated portfolios in the pool.

Rescaling to a Target Volatility

When using unit allocations, such as the number of futures contracts rather than a percentage weighting factor, it is necessary to rescale each portfolio allocation to a target volatility, or risk level. Trading too few contracts will yield a low risk and a low return. It is necessary to specify the risk as a probability of loss equal to a percentage of the investment. If the target volatility is 15%, and the risk evaluated after the random allocation shows risk of only 5%, then all allocations should be multiplied by 3. Portfolio specifications will allow individual portfolio risk to vary from the targeted risk by ±5% (specified by the user). After rescaling, it will be necessary to see if any of the assets exceed its

TABLE 24.7 Evaluating Portfolio Return and Risk

Step 4. Evaluate the Returns of Portfolio 1

| | Weighting Factors | | | | | | |
| | 0.131 | 0.290 | 0.061 | 0.310 | 0.208 | | |
	A	B	C	D	E	Portfolio NAV	Returns
8/23/2000	100.00	100.00	100.00	100.00	100.00	100.00	
8/24/2000	100.00	99.24	99.09	98.83	99.15	99.18	−0.0082
8/25/2000	99.93	100.02	100.02	100.03	100.02	100.01	0.0083
8/28/2000	101.66	101.66	101.98	102.55	101.86	102.00	0.0199
8/29/2000	101.39	101.40	101.33	102.16	101.58	101.67	−0.0032
8/30/2000	100.35	102.54	101.25	103.90	102.85	102.66	0.0098
8/31/2000	99.94	102.15	101.06	103.30	103.33	102.40	−0.0026
9/1/2000	100.47	101.58	101.06	102.43	103.33	102.03	−0.0036
9/5/2000	101.40	102.00	101.56	102.20	103.80	102.33	0.0029
9/6/2000	102.45	104.09	103.82	105.41	105.26	104.51	0.0213
9/7/2000	102.45	104.67	103.82	106.86	105.26	105.13	0.0059
9/8/2000	102.45	104.67	103.82	101.40	105.26	103.44	−0.0161
9/11/2000	101.22	104.05	103.88	100.52	104.72	102.71	−0.0070
9/12/2000	99.45	102.38	103.88	97.94	102.84	100.81	−0.0186
9/13/2000	98.50	101.48	103.88	96.56	101.83	99.78	−0.0102
9/14/2000	99.01	100.66	103.88	95.60	100.91	99.12	−0.0066
9/15/2000	102.83	101.45	105.03	96.30	100.91	100.14	0.0103
9/18/2000	104.48	103.07	106.63	99.30	100.91	101.85	0.0171
					1.5315	0.2933	0.1915
					Ratio	Return	Risk

liquidity limits. If yes, the portfolio is discarded. Note that the rescaling is not needed to find the best portfolio, only to satisfy constraints, because rescaling does not change the information ratio.

Test for Completion

The objective function has now been calculated for each portfolio in the pool. The best objective function in the pool is now compared to the best objective function in the previous cycle to decide if the process has been completed. This step is skipped on the first iteration, following the initialization of the pool. Completion may be tested in two ways:

1. *Convergence.* The best objective function in the current pool is identical, or nearly identical, to the best objective function in the previous n iterations. It is best to allow the method 2 or 3 cycles without any appreciable improvement before ending the process. It may be better if the best 5 returns are unchanged.

2. *Number of iterations.* The process has completed a fixed number of iterations.

Although convergence is the normal completion criterion in search methods, a fixed number of iterations may serve you best in this method. The genetic algorithm technique rapidly approaches an optimal solution, but it may show very slow improvement afterwards. Because this level of accuracy is not necessary, it is more important to repeat the entire process 3 to 5 times beginning with a new pool of random strings. Similar results from these tests, either from actually weighting factors or the objective function, will confirm that the method has found a global maximum rather than a local maximum. This concept will be discussed further.

Propagation

Darwin observed that propagation of the species was improved by the *survival of the fittest*. This step in the genetic algorithm process replicates the survival of the fittest. The purpose of this propagation step is to allow strings with highest objective functions to survive the selection process while those with the lowest objective functions are discarded.

Ranking the Portfolios

Using the objective functions, the portfolios in the pool are sorted so that the ones with the highest objective functions are at the top. For implementation, it is faster to do an *index sort*. First, number the portfolios sequentially by creating a column of the portfolio numbers alongside their respective objective functions, then sort only those two columns. This avoids shuffling large lists of assets, which can be time-consuming, even for fast computers.

Another random process is used to propagate the best portfolios. If there are only 8 strings in the pool with the following sorted objective functions:

$$3.65, 2.50, 2.00, 1.65, 1.40, 1.20, 1.05, 1.00$$

then the differences between these values will be used to encourage the reproduction of the portfolios with the best objective functions. The differences are:

$$1.15, 0.50, 0.35, 0.25, 0.20, 0.15, 0.05,$$

representing a total span of $3.65 - 1.00 = 2.65$. These differences, converted to percentages of the span, are:

$$43.4, 18.9, 13.2, 7.5, 5.6, 1.9$$

From this it can be seen that there are much larger gaps between the higher objective functions than between the smaller ones. This is the feature used to facilitate propagation. For different objective functions, with different ranges of results, unique propagation rules will be needed.

Propagation of the Best Results

Propagation enables the best portfolios to not only survive, but have added influence on the outcome.[6] Propagation, along with most of the other steps in the genetic algorithm solution, uses random numbers. In this case, a *uniform random number* between 3.65 and 1.00, the highest and lowest objective values, is needed. It is important to generate a uniform random number, one that is equally distributed within the range of values. To do this, the Excel function *randbetween(lowest,highest)* will be used, where *lowest* = 100 and *highest* = 365 (the function only uses integers; therefore the values are multiplied by 100).

If there are 8 portfolios in the sample pool, we first find the midpoint between each of the 8 objective functions:

Portfolio	1	2	3	4	5	6	7	8
Objective functions	3.65	2.50	2.00	1.65	1.40	1.20	1.05	1.00
Midpoints		3.075	2.25	1.825	1.525	1.30	1.125	1.025

Next, generate 8 random numbers, one for each portfolio. Those portfolios chosen to be propagated are placed into a new pool of the same size as the original. Although the random numbers are between 100 and 365, they are divided by 100 to fit the scale of the objective functions. When the random number falls to the left of the midpoint of two objective functions, one copy of the portfolio *to the left* is placed in the new pool, as shown below:

Random Number	Reason	No. of Copies
3.63	Closer to 1 than 2	1 copy of 1
3.49	Closer to 1 than 2	1 copy of 1
1.71	Closer to 4 than 3	1 copy of 4
1.51	Closer to 5 than 4	1 copy of 5
1.93	Closer to 3 than 4	1 copy of 3
1.44	Closer to 5 than 6	1 copy of 5
2.64	Closer to 2 than 1	1 copy of 2
3.48	Closer to 1 than 2	1 copy of 1

When completed, the method has created a new pool containing the following 8 portfolios (shown in order of highest objective function):

New portfolio	1	2	3	4	5	6	7	8
Original portfolio	1	1	1	2	3	4	5	5
Objective functions	3.65	3.65	3.65	2.50	2.00	1.65	1.40	1.40

---

[6]Thanks to Lutz Borgiel of ARIAD Asset Management in Hamburg, Germany, for his help with this propagation method.

Because the original portfolio 1 was the fittest, with an objective function of 3.65, it has the greatest chance of surviving. Portfolios 6, 7, and 8 did not survive, and while portfolio 5 has more copies than the better results 2 through 4, it will eventually be eliminated unless it mutates into a better portfolio. When performing this process with a large pool of 5,000 or 100,000 strings, many of the old portfolios will survive. The objective is to remove about 10% of the worst performers on each cycle while propagating the best.

The following steps define the propagation process for a large pool of P portfolios. As individual portfolios (p) are propagated, they are moved from the original pool 1 to a new pool 2:

1. Calculate the objective function for all P portfolios in the pool.
2. Sort the portfolios, putting the highest objective function in position 1.
3. Note the range of the objective functions from the highest OF_H in position 1 to the lowest OF_L in position p.
4. Get a uniform random number, R, between OF_L and OF_H.
5. Determine which portfolio, p, has the objective function closest to the random number R.
6. Make 1 copy of portfolio p and place it into the new pool 2.
7. If P portfolios have been propagated, then go to Step 8; otherwise, go to Step 4.
8. Erase the contents of pool 1 and copy the contents of pool 2 into pool 1.

Pool 1 now contains the propagated portfolios. Pool 2 is only used as temporary storage during the propagation process.

If the propagation process is too slow, it can be accelerated by multiplying the number of copies by a factor, such as 2, or by using the integer value of the objective function as the number of copies. The danger is that the program will make too many copies of the same strings, reducing the diversity in the gene pool and increasing the chance that the final choice will come from the original pool. This is the equivalent of finding a local maximum rather than a global one.

Mating Parents to Create Offspring

Mating is the process of combining two *parent* (original) strings in a way that creates two unique offspring. As in life, if two strings with excellent genes are combined, the offspring may have genes that are superior to either of the parents. This step will also be implemented using random numbers.

The mating process will select two portfolios at random from the pool of P portfolios and combine part of one string with part of the other. The strings will first be divided at a random point, then the two left parts will be switched. The result is two unique portfolios. Mating follows the steps:

1. *Select portfolio p_i* using a random whole number between 1 and P, where P is the total number of portfolios. Mark portfolio i as processed so that it will not be selected again.

2. *Select portfolio p_j* using another random number between 1 and P. If the portfolio selected is marked as processed, then increase the random number by 1 until the next unused portfolio is found. If the value reaches P, then begin again at 1.

3. *Get a random number*, a, from 1 to the number of assets in the portfolio. Our example uses 5 assets; therefore the random number must be an integer from 1 to 5.

4. *Mate* the two portfolios by switching the weighting factors in p_i, assets 1 through a, with the weighting factors in p_j, assets 1 through a. For example, if $a = 2$, then the two leftmost assets of portfolios i and j are switched:

Before mating (*parents*)

Portfolio i:	w_{i1}	w_{i2}	w_{i3}	w_{i4}	w_{i5}
Portfolio j:	w_{j1}	w_{j2}	w_{j3}	w_{j4}	w_{j5}

After mating (*offspring*)

Portfolio i:	w_{j1}	w_{j2}	w_{i3}	w_{i4}	w_{i5}
Portfolio j:	w_{i1}	w_{i2}	w_{j3}	w_{j4}	w_{j5}

5. *Normalize the two offspring.* Because the sum of the weighting factors will no longer total 1 in each of the offspring portfolios, they must be normalized by dividing each weighting factor by the sum of all the weighting factors in the new portfolio.

6. *Correct for liquidity constraints.* Mating does not affect individual market liquidity constraints because the positions in the portfolio of those items being switched remain the same. Group liquidity, that is, a maximum volume that can be traded in a set of markets, can be affected if the split occurs within one of these groups. If the petroleum sector—crude oil, heating oil, and unleaded gasoline—is considered a futures market group, then each market may be limited to 2,000 contracts, but the total group may also be limited to 4,000 contracts because of diversification. During the initial allocation of the pool, the allocations for the three energy markets are reduced to 4,000 if they exceed that amount. For example, if crude, heat, and gas have initial allocations of 2,100, 1,500, and 1,500, they could be scaled to 1,400, 1,000, and 1,000 to keep the total below the group maximum of 4,000. However, another portfolio allocation might have 800, 1,500, and 1,500. If these two portfolios are mated and the split comes after the first market, one of the offspring will have allocations of 1,400, 1,500, and 1,500, which exceeds the limit by 400. Therefore, groups need to be rescaled to satisfy liquidity constraints after mating.

7. *Continue mating* until all the portfolios have been combined into offspring. It may seem as though the act of mating all the parents might lose the best combination; therefore only 10% to 20% should be mated. The propagation of duplicate strings and the large number of combinations that might yield excellent results favor the mating of all portfolios in the pool. Less than that will cause convergence to an early, suboptimal solution. It is convenient to be able to specify the percentage of strings to be mated.

The mating process creates new combinations of portfolio allocations. Because of the previous propagation step, there are more copies of portfolios that generated higher objective functions. By combining these portfolios randomly, the chance of finding a better combination is very high.

Mutating: Introducing New Genes into the Pool

The final step in Darwin's process is mutation, a random change that can improve the species or create an inferior member. The act of mutating prevents the gene pool from stagnating. In this application, introducing new genes prevents the selection process from converging too soon. Consistent with the other steps, mutation replaces one randomly selected gene (one weighting factor) in a randomly selected portfolio, with a new, random value, in about 10% of the current portfolios.

1. Get a random number between 1 and P to indicate the portfolio to be mutated. Mark that portfolio as *processed* so that it does not receive another mutation.
2. Get a random number between 1 and a, the number of assets. This will be the asset to be changed.
3. Get a random number between 0 and 1.0 to represent the new weighting factor.
4. Adjust the new weight to satisfy any individual market liquidity constraints and group constraints (as discussed in the previous section, "Mating," item 6).
5. Normalize the weighting factors of the mutated portfolio by dividing each weighting factor by the sum of the weighting factors being sure that the constraints are still satisfied.
6. Continue mutating portfolios until the targeted percentage, between 5% and 15%, is reached.

As in the other steps, too much mutation will result in losing the best results; too little mutation will cause the genetic algorithm to converge too quickly to one of the portfolios already in the pool.

Continue with the Evaluation Step

The genetic algorithm process is complete. The pool has been created, propagated, mated, and mutated. The results can now be reevaluated by calculating the objective function of each of the portfolios to see if there has been an improvement. This is done in the previous section "Evaluation of Each String in the Pool." The process continues until the completion criterion is satisfied.

Optimal Subsets

When solving the portfolio problem for a large number of assets, for example 50 to 100, it may be practical to find the best subset of only 15 assets. While the genetic algorithm

process will reject systems that do not contribute to the solution, the best allocations may concentrate 90% of the allocations in 10% of the assets. The remaining assets provide only marginal improvement.

The portfolio solution will normally allocate some amount of the investment to all assets, unless those assets are nearly identical to another asset in the portfolio; therefore, any subset will be suboptimal. Using subsets is contrary to the goals of the Markowitz solution because it limits the diversification. The mean-variance solution will always seek the lowest variance, which comes with greater diversification. Nevertheless, a portfolio of the most important assets is practical for an investor.

One solution to identifying the best subset is to run the full program on all assets, then simply discard the ones with small allocations and redistribute the investment in proportion to the remaining allocations. This approach is unnecessarily time-consuming and may yield allocations that exceed liquidity constraints when the allocations are redistributed to the chosen subset.

Using the genetic algorithm, subsets can be established during the initialization of the pool and maintained throughout the rest of the process. For a subset of s assets out of a total of S assets,

1. When initializing the pool, choose only s assets at random. The remaining assets are given the weight of zero. Follow the same steps for choosing weighting factors, then normalize.

2. Propagation remains the same. Only the objective function is calculated; the number of assets is not relevant.

3. Mating requires an additional step. Select the first parent portfolio and decide where it is to be split. Count the number of assets on each side of the splitting point. Select another parent portfolio. Count the number of assets on each side of the same splitting point. If they are the same, those two portfolios can be mated by switching assets at the splitting point. If not, move to the next sequential portfolio and test if the number of assets match. Continue until a match is found. Two parents cannot be mated if they have a different number of assets on each side of the split because the resulting offspring will have either more or less assets than the required subset amount s. Except for a few portfolios, the normal distribution of assigned assets due to random selection will allow most portfolios to be matched with others that have the same asset distribution relative to the split. The few portfolios that do not match can be discarded.

4. Mutating has a problem similar to mating, but it is easier to solve. After selecting a portfolio at random, choose an active asset at random and remove it. Then choose another asset (previously zero but not the one just removed) and generate the new weight using a random number, normalizing afterward. This keeps the total number of assets at s.

The subset solution is much faster than a full solution. It involves some compromise because it is a suboptimal solution. Without controlling the mating and mutating process,

the program will try to introduce more assets in order to improve the results. Although the pool starts with portfolios that all have only *s* assets, the mating and mutating process will continue to combine, add, and propagate portfolios with more assets. You will be fighting nature, although the best solution using a large number of assets will be only slightly better than the optimal subset.

Is It the Best Portfolio?

For a small number of assets, the result of GASP will be the optimal portfolio. As the number of assets gets larger, the number of possible solutions quickly becomes too large to have included all combinations in the pool, but for 20 or even 50 assets the solution is very likely the optimal combination. For a much larger number of assets, the combinations created by the genetic algorithm in a limited amount of computing time may be a small percentage of the total possible. Yet the answer may be close to optimal. At least, it will be very good.

When allocating many assets, one way of estimating the answer is to limit the percentage increments assigned to each allocation. Using 100 assets as an example, a random generation of weighting factors could return values with 4 decimal places, such as 0.0523, representing a 5.23% allocation. Then each of the 100 assets would have 10,000 possible choices and the total combinations would be 10,000! $(1 \times 2 \times 3 \times \cdots \times 10,000)$, an extraordinarily large number. However, if the problem was limited to whole percent, the number of combinations would only be 100!, still very large but much more manageable. When arriving at the optimal solution, the shape of the allocations based on 1% increments must be similar to that based on 0.001% increments. Except for a very large investment, an allocation of 0.001% of the investment is not significant.

Although there are many answers with near-optimal solutions, many of them differ in the second and third decimal place, which makes it unimportant to the trader. The GASP program should consistently return a near-optimal solution.

Verifying the Global Maximum

One way of verifying that the solution is a global maximum and not a local one is to run the solution more than once. Each new pass creates a completely new pool of random allocations and begins the genetic algorithm process again. If five passes were made on a portfolio problem with 50 assets, the allocations should be very similar. The one with the highest objective function is the best, and may represent the global maximum. There may also be many local maxima with peaks that are nearly the same. This can be seen if the objective functions are the same but the allocations are significantly different.

It is more efficient to run the portfolio solution 5 times, each with 50 to 100 cycles of the genetic algorithm, than to run it once with 500 cycles. The solution begins to converge after 50 cycles and is unlikely to change significantly afterwards. This is due to the percentage of mating and mutation specified in the program at 10% to 15%. If a larger percentage is used, then convergence would not be as fast, but the results could become unstable.

An Example of a GASP Solution

The best way to understand the full process of the genetic algorithm solution is to see the step-by-step results. This example uses 19 unique strategies for trading the NASDAQ 100 futures contract, each with three components that could be based on a trend (*T*), mean-reversion (*M*), or a combination of both *T* and *M* (*B*). These strategies are applied to only long positions (*L*) or only short sales (*S*). Therefore, a long-only strategy that combines trend, mean-reversion, and mean-reversion components will be labeled LTMM, and appears as the eighth strategy in the output, shown in Figure 24.7. Other features of this test are:

1. A total of 19 assets (in this case an *asset* is the return stream of one NASDAQ strategy).
2. The historic returns of each strategy cover the interval from May 4, 1995, through June 30, 2002, a span of 7.16 years.
3. A target optimal subset of 12 of 19 assets.
4. An investment of $10 million.
5. An initial pool size of 5,000.
6. A mutation rate of 10%.
7. A risk calculation based on semivariance of drawdowns (*D*).
8. A target risk of 16.7% representing the confidence level of 99.74.
9. Each pass of the program will have 100 cycles.
10. There will be 3 passes.

Allocations that appear in Figure 24.7 are in numbers of futures contracts. The program knows the margin needed to trade one contract and the margin-to-equity ratio required by the investor. The total number of contracts allocated by the GASP program is a function of the total investment and the margin-to-equity ratio.

Overview of the Process

The evolution of results in Figure 24.7 can be seen in the number of cases (*Case*) and the objective function (*ObjF*) in the first two columns, and in the 19 columns of contract allocations under the strategy names, representing the largest part of the printout. In Pass 1, Case 1, at the top of part *a*, the objective function is 12.955. The allocations represent the best objective function found in the initial pool of 5,000 that qualified as having a risk within 5% of the targeted risk level of 2.85 using the semivariance. Note that all values in the column *StdD* are within 5% of 2.85. Allocations are always rescaled, if possible, to achieve a standard risk level in order to avoid low-risk solutions that also have low returns. The columns that follow *ObjF* are alternate risk measures:

Std is the standard deviation of the daily returns.

StdR is the standard deviation of the residuals, based on a linear regression of the return series.

(a) Pass 1

GENETIC ALGORITHM foliondx   20020705 9:06:57

from 19950504 to 20020630, 7.16 years, 19 assets, investment $10 mil
passes 3, cases 100, subset 12, pool size 5000, cull 2.00, mutate 0.10 preset fixed
basis D, minliq 1, liq on, ratio on, riskeq off, maxsys 0, %target 16.97 16.60 6.67, adjtarget 2.85, VARlimit 5%

PASS 1

Case	ObjF	Std	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMB	LBMB	LMMB	LTMM	LBMM	LMMM	STBB	SBBB	STBM	SBBM	STMB	SBMB	STMM	SBMM	SMMM	Net
1	12.955	1.18	19.17	2.79	16.85	0	0	0	1	19	40	41	11	21	53	9	71	0	18	0	0	27	35	0	346
4	13.089	1.11	15.17	2.75	19.42	0	0	0	39	15	33	34	9	17	44	7	28	0	33	0	17	34	0	0	310
5	13.905	1.26	15.00	2.78	16.85	5	0	17	35	0	42	42	0	18	49	8	66	0	16	0	0	25	14	0	337
9	14.161	1.41	14.21	2.85	14.75	7	0	32	26	2	25	117	0	4	39	42	2	0	26	0	0	33	0	0	355
10	14.529	1.43	15.86	2.87	16.84	0	20	27	41	2	25	117	0	4	39	42	2	0	26	0	0	33	0	0	378
12	14.944	1.36	13.94	2.72	16.02	13	0	15	26	0	29	110	0	3	36	6	55	0	29	0	0	33	0	0	352
13	15.111	1.39	17.16	2.79	16.25	13	0	15	26	0	29	110	0	3	36	6	55	0	29	0	0	33	0	0	376
14	15.439	1.48	17.98	2.89	16.37	0	0	7	29	23	35	117	0	4	44	5	65	14	16	0	0	25	24	0	404
15	15.925	1.50	19.22	2.92	17.72	0	3	32	32	0	45	117	0	0	39	2	71	0	0	0	34	33	0	0	412
18	16.293	1.47	18.61	2.78	15.65	5	0	7	39	15	25	117	0	4	45	7	61	0	14	0	20	33	0	0	407
19	16.313	1.48	18.89	2.79	12.90	0	3	32	35	0	31	117	0	3	44	7	71	0	32	0	39	33	0	0	416
22	16.477	1.48	17.18	2.79	14.01	5	0	27	41	0	33	109	0	3	43	7	71	0	16	0	20	31	0	0	406
25	16.732	1.46	19.12	2.71	15.14	0	0	17	39	15	25	117	0	0	44	7	65	0	16	0	29	31	6	0	411
29	16.810	1.47	18.23	2.71	13.58	3	0	31	35	0	26	109	0	3	49	7	71	0	16	0	27	33	0	0	410
30	16.833	1.47	20.29	2.73	14.33	0	0	17	39	15	25	117	0	0	44	7	71	0	32	0	23	33	0	0	423
31	16.842	1.48	18.58	2.78	14.70	0	0	33	38	0	40	103	0	3	49	8	66	0	20	0	23	33	0	0	416
32	16.972	1.50	18.17	2.72	13.37	3	0	31	35	0	31	117	0	2	45	7	71	0	20	0	19	33	0	0	414
34	17.044	1.45	17.45	2.71	14.68	4	0	29	41	0	37	96	0	0	49	8	66	0	15	0	26	33	0	0	404
36	17.078	1.46	17.61	2.72	14.71	4	0	29	41	0	37	96	0	0	49	7	67	0	19	0	24	33	0	0	406
37	17.100	1.49	17.77	2.75	14.91	5	0	27	41	0	38	102	0	0	49	7	59	0	28	0	23	33	0	0	412
39	17.122	1.48	18.00	2.73	14.85	3	0	27	41	0	38	102	0	3	49	8	66	0	15	0	26	33	0	0	411
41	17.129	1.48	17.09	2.72	14.38	4	0	33	41	0	29	102	0	3	49	8	66	0	15	0	24	33	0	0	407
42	17.152	1.50	18.42	2.76	14.86	4	0	27	41	0	40	103	0	0	49	8	71	0	20	0	23	33	0	0	419
43	17.190	1.48	18.13	2.73	14.70	4	0	27	41	0	38	102	0	0	49	8	66	0	16	0	29	33	0	0	413
44	17.225	1.48	18.96	2.73	14.46	3	0	29	41	0	40	108	0	0	44	14	71	0	16	0	22	33	0	0	421
45	17.235	1.48	18.01	2.71	14.70	4	0	27	41	0	38	102	0	0	49	8	66	0	20	0	24	33	0	0	412
47	17.257	1.48	18.53	2.73	14.68	3	0	27	41	0	40	103	0	0	49	8	71	0	20	0	22	33	0	0	417
49	17.261	1.48	18.10	2.71	14.68	3	0	27	41	0	40	103	0	0	49	8	66	0	18	0	24	33	0	0	412
52	17.285	1.50	18.43	2.74	14.41	3	0	31	41	0	37	107	0	0	49	14	66	0	16	0	24	33	0	0	421
54	17.306	1.48	18.17	2.72	14.39	3	0	33	41	0	36	102	0	0	49	14	66	0	20	0	20	33	0	0	417
59	17.315	1.48	17.98	2.71	14.39	3	0	33	41	0	36	102	0	0	49	14	66	0	16	0	22	33	0	0	415

FIGURE 24.7 GASP solution for 16 NASDAQ 100 futures strategies.

Case	ObjF	Std	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMM	LBMM	LMMB	LMMM	STBB	SBBB	STBM	SBBM	STMM	SBMM	STMB	SBMB	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
objf	17.315	1.48	17.98	2.71	14.39	3	0	33	41	0	36	102	0	0	49	14	66	0	16	0	22	33	415
Rtio	17.315	1.48	17.98	2.71	14.39	3	0	33	41	0	36	102	0	0	49	14	66	0	16	0	22	33	415

BEST RATIO SOLUTION

Norm	AROR	AStd	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMM	LBMM	LMMB	LMMM	STBB	SBBB	STBM	SBBM	STMM	SBMM	STMB	SBMB	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
stdev	33.29	16.60	12.75	1.92	10.20	2	0	23	29	0	26	72	0	0	35	10	47	0	11	0	16	23	294
sdres	17.42	0.55	6.67	1.01	5.34	1	0	12	15	0	13	38	0	0	18	5	24	0	6	0	8	12	152
maxdd	65.26	2.05	24.98	3.77	20.00	4	0	46	57	0	50	142	0	0	68	19	92	0	22	0	31	46	577

EXPONENTIAL DISTRIBUTION CONFIDENCE LEVEL 16.97% EQUAL TO MAXIMUM DRAWDOWN OF $ 1.7 million

Conf	AROR	Std	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMM	LBMM	LMMB	LMMM	STBB	SBBB	STBM	SBBM	STMM	SBMM	STMB	SBMB	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
95.40	95.38	3.00	36.54	5.51	29.28	6	0	67	83	0	73	207	0	0	100	28	134	0	33	0	45	67	843
99.74	49.46	1.55	18.97	2.87	15.19	3	0	35	43	0	38	107	0	0	52	15	69	0	17	0	23	35	437
99.90	42.59	1.34	16.26	2.47	13.05	3	0	30	37	0	33	92	0	0	44	13	60	0	14	0	20	30	376

Best case has 11 of 19 assets.

FIGURE 24.7 *(Continued)*

(b) Pass 2

Case	ObjF	Std	StdR	StdD	Mdd%	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
						LTBB	LBBB	LTBM	LBBM	LTMB	LBMB	LMMB	LTMM	LBMM	LMMM	STBB	SBBB	STBM	SBBM	STMB	SBMB	STMM	SBMM	SMMM	Net
1	13.651	1.18	13.66	2.75	16.61	14	0	12	49	0	3	34	4	1	58	0	0	3	38	0	0	53	22	0	291
2	14.190	1.19	12.82	2.73	16.58	12	0	13	53	0	4	37	5	2	63	0	0	4	41	0	0	57	3	0	294
5	14.449	1.44	21.58	2.82	18.82	10	0	6	36	0	0	123	0	0	62	26	44	2	38	19	9	28	0	0	403
9	14.576	1.37	16.40	2.82	15.56	0	0	34	18	0	29	105	7	0	55	26	8	0	0	13	37	31	0	0	363
10	14.595	1.39	17.56	2.91	18.90	0	0	0	54	9	29	105	7	0	55	26	8	0	0	13	37	31	0	0	374
11	15.225	1.34	19.08	2.73	15.85	0	0	33	17	0	28	101	0	12	50	18	26	0	30	0	0	43	17	0	375
13	15.246	1.42	19.73	2.93	17.95	0	0	41	22	0	28	101	0	12	50	18	26	1	32	0	0	53	11	0	392
14	16.199	1.45	22.19	2.84	14.48	0	0	35	30	0	29	99	0	0	62	30	44	1	27	0	0	43	18	0	425
17	16.387	1.41	20.51	2.75	17.48	6	0	17	41	0	27	99	0	0	62	18	44	0	27	0	0	40	19	0	412
18	16.561	1.45	21.51	2.81	15.30	0	0	35	30	0	29	101	0	0	62	18	53	2	32	0	0	49	10	0	421
21	16.669	1.44	20.76	2.80	15.63	0	0	30	42	0	38	95	0	0	58	24	41	1	29	0	0	39	22	0	419
22	16.798	1.45	19.27	2.78	15.66	0	0	31	43	0	29	105	7	0	55	26	36	0	22	0	13	37	14	0	418
23	16.980	1.48	19.02	2.83	16.68	0	0	26	48	0	41	105	7	0	55	26	36	0	33	0	13	37	0	0	427
25	17.013	1.52	21.67	2.90	16.56	0	0	31	43	0	41	101	0	0	62	26	37	0	28	0	17	40	16	0	442
26	17.077	1.51	21.10	2.87	16.27	0	0	31	43	0	41	101	0	0	62	26	37	0	28	0	17	40	12	0	438
27	17.198	1.44	18.91	2.73	16.01	0	0	30	42	0	36	94	4	0	62	23	44	1	33	0	13	37	0	0	419
29	17.299	1.45	18.97	2.73	15.93	0	0	30	42	0	35	97	4	0	62	23	44	1	33	0	13	37	0	0	421
31	17.427	1.46	19.25	2.73	15.97	0	0	31	42	0	38	95	0	0	62	18	53	1	23	0	22	37	0	0	422
35	17.479	1.46	19.51	2.72	15.65	0	0	29	46	0	32	101	0	0	62	18	41	0	27	0	16	41	6	0	419
38	17.651	1.51	19.24	2.74	15.77	0	0	30	48	0	32	105	0	0	62	18	44	1	29	0	22	37	0	0	428
42	17.661	1.50	19.55	2.75	15.73	0	0	32	50	0	29	101	0	0	62	17	53	2	27	0	22	37	0	0	432
45	17.677	1.50	19.42	2.74	15.79	0	0	32	47	0	32	101	0	0	62	18	53	1	25	0	22	37	0	0	430
47	17.694	1.50	19.01	2.72	15.64	0	0	30	50	0	29	105	0	0	62	18	41	1	30	0	22	37	0	0	425
48	17.710	1.49	18.87	2.71	15.70	0	0	32	50	0	29	101	0	0	62	18	44	2	27	0	22	37	0	0	424
51	17.730	1.51	18.89	2.73	15.97	0	0	30	47	0	30	101	0	0	62	18	44	0	26	0	22	37	0	0	426
53	17.767	1.51	19.35	2.72	15.88	0	0	29	46	0	35	105	0	0	62	18	44	1	29	0	22	37	0	0	428
56	17.800	1.51	19.49	2.75	15.73	0	0	32	50	0	35	105	0	0	62	17	44	0	30	0	22	40	0	0	431
72	17.801	1.52	18.68	2.72	15.81	0	0	32	47	0	32	107	0	0	62	17	44	0	30	0	20	36	0	0	427

FIGURE 24.7 (Continued)

Case	ObjF	Std	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMB	LBMB	LMMB	LTMM	LBMM	LMMM	STBB	SBBB	STBM	SBBM	STMB	SBMB	STMM	SBMM	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
objf	17.801	1.52	18.68	2.72	15.81	0	0	32	47	0	32	107	0	0	62	17	44	0	30	0	20	36	0	0	427
Rtio	17.801	1.52	18.68	2.72	15.81	0	0	32	47	0	32	107	0	0	62	17	44	0	30	0	20	36	0	0	427

BEST RATIO SOLUTION

Norm	AROR	AStd	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMB	LBMB	LMMB	LTMM	LBMM	LMMM	STBB	SBBB	STBM	SBBM	STMB	SBMB	STMM	SBMM	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
stdev	33.28	16.60	12.81	1.87	10.85	0	0	22	32	0	22	73	0	0	43	12	30	0	21	0	14	25	0	0	294
sdres	17.32	0.54	6.67	0.97	5.65	0	0	11	17	0	11	38	0	0	22	6	16	0	11	0	7	13	0	0	152
maxdd	61.35	1.93	23.62	3.45	20.00	0	0	40	59	0	40	135	0	0	78	22	56	0	38	0	25	46	0	0	539

EXPONENTIAL DISTRIBUTION CONFIDENCE LEVEL 16.97% EQUAL TO MAXIMUM DRAWDOWN OF $ 1.7 million

Conf	AROR	Std	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMB	LBMB	LMMB	LTMM	LBMM	LMMM	STBB	SBBB	STBM	SBBM	STMB	SBMB	STMM	SBMM	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
95.40	98.10	3.08	37.74	5.51	31.96	0	0	65	95	0	65	216	0	0	125	34	89	0	61	0	40	73	0	0	863
99.74	50.60	1.59	19.60	2.94	16.49	0	0	33	49	0	33	112	0	0	65	18	46	0	31	0	21	38	0	0	446
99.90	43.65	1.37	16.72	2.46	14.28	0	0	29	42	0	29	96	0	0	56	15	40	0	27	0	18	32	0	0	384

Best case has 11 of 19 assets.

FIGURE 24.7 (Continued)

(c) Pass 3

PASS: 3

Case	ObjF	Std	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMB	LBMB	LMMB	LTMM	LBMM	LMMM	STBB	SBBB	STBM	SBBM	STMB	SBMB	STMM	SBMM	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
1	13.662	1.20	17.41	2.73	13.02	0	0	38	30	0	12	47	0	18	48	0	40	0	40	14	26	27	0	0	340
3	13.691	1.24	17.36	2.82	13.64	0	6	38	34	0	12	47	0	18	48	0	40	0	40	14	26	27	0	0	350
4	13.802	1.24	19.10	2.80	14.51	0	0	31	42	0	12	47	0	18	48	0	40	0	40	14	23	28	10	0	353
5	13.936	1.22	18.54	2.83	17.35	0	0	0	66	17	17	36	10	0	48	0	40	0	40	14	23	28	10	0	349
7	14.292	1.18	17.14	2.72	16.43	0	0	0	66	17	17	36	10	0	48	0	40	0	40	14	23	28	0	0	339
9	14.731	1.45	16.53	2.89	15.25	0	18	34	41	0	18	110	0	28	41	30	21	3	28	13	0	19	0	0	404
10	15.308	1.38	17.13	2.74	12.94	0	18	34	34	0	16	102	0	0	44	0	59	0	37	14	0	25	0	0	383
12	15.484	1.40	18.65	2.76	13.53	0	0	22	28	9	38	102	0	0	44	0	37	0	40	10	26	27	12	0	395
13	15.776	1.36	18.85	2.72	16.68	0	0	18	23	11	38	100	0	3	48	0	40	0	40	14	10	35	0	0	380
15	16.432	1.40	20.95	2.72	15.11	0	0	23	31	6	18	96	0	0	55	12	45	0	40	13	24	27	0	0	413
18	16.621	1.47	18.55	2.79	15.45	0	0	23	58	6	38	106	0	0	51	9	35	0	37	0	25	25	0	0	412
20	16.941	1.48	18.99	2.76	14.30	0	0	25	49	0	18	115	0	0	48	22	42	0	28	10	27	33	0	0	422
22	17.063	1.49	21.29	2.80	16.56	0	0	20	58	6	18	107	0	0	55	12	45	0	30	0	21	35	0	0	431
24	17.147	1.47	19.75	2.72	13.52	0	0	27	58	17	16	114	0	0	48	22	45	0	28	0	26	33	0	0	427
25	17.225	1.50	18.93	2.81	16.44	0	0	25	60	6	20	107	0	0	51	11	54	0	37	4	27	33	0	0	426
26	17.370	1.49	18.42	2.71	14.19	0	0	27	58	6	25	106	0	0	55	10	42	0	37	0	25	32	0	0	416
27	17.459	1.54	19.59	2.80	15.34	0	0	27	58	6	22	115	0	0	55	12	39	0	32	5	25	38	0	0	433
30	17.554	1.52	19.40	2.77	14.68	0	0	27	58	6	20	107	0	0	55	20	45	0	30	5	25	32	0	0	437
33	17.587	1.52	19.98	2.78	15.34	0	0	27	58	6	18	110	0	0	54	12	45	0	36	5	23	35	0	0	435
34	17.600	1.55	20.44	2.81	14.93	0	0	25	62	6	20	115	0	0	51	11	54	0	36	5	25	33	0	0	445
35	17.736	1.52	20.17	2.73	14.33	0	0	25	58	6	20	109	0	0	55	11	54	0	26	3	25	33	0	0	437
40	17.737	1.53	19.53	2.74	14.85	0	0	27	58	6	18	115	0	0	55	12	54	0	26	5	25	33	0	0	434
41	17.758	1.54	19.80	2.77	14.94	0	0	25	58	6	20	115	0	0	55	11	54	0	26	5	25	35	0	0	438
42	17.760	1.51	19.90	2.72	14.33	0	0	23	58	6	21	107	0	0	55	12	54	0	37	3	22	35	0	0	433
43	17.848	1.52	19.85	2.72	14.81	0	0	25	58	6	20	113	0	0	55	11	54	0	26	5	25	35	0	0	434
51	17.848	1.53	20.03	2.72	14.93	0	0	25	58	6	21	115	0	0	55	12	54	0	26	5	25	35	0	0	435
52	17.859	1.53	19.90	2.73	14.86	0	0	25	58	6	20	115	0	0	55	12	54	0	26	5	25	35	0	0	436
55	17.861	1.53	20.01	2.74	14.89	0	0	25	58	6	20	115	0	0	55	12	54	0	26	5	26	35	0	0	437
57	17.864	1.53	19.90	2.73	14.97	0	0	25	58	6	20	115	0	0	54	11	54	0	26	5	25	35	0	0	435
59	17.870	1.52	19.83	2.71	14.66	0	0	25	58	6	19	113	0	0	55	12	54	0	26	5	25	35	0	0	433
62	17.894	1.52	19.88	2.71	14.82	0	0	25	58	6	19	115	0	0	54	11	54	0	26	5	26	35	0	0	434

FIGURE 24.7 *(Continued)*

Case	ObjF	Std	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMB	LBMB	LMMB	LTMM	LBMM	LMMM	STBB	SBBB	STBM	SBBM	STMB	SBMB	STMM	SBMM	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
objf	17.894	1.52	19.88	2.71	14.82	0	0	25	58	6	19	115	0	0	54	11	54	0	26	5	26	35	0	0	434
Rtio	17.894	1.52	19.88	2.71	14.82	0	0	25	58	6	19	115	0	0	54	11	54	0	26	5	26	35	0	0	434

BEST RATIO SOLUTION

Norm	AROR	AStd	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMB	LBMB	LMMB	LTMM	LBMM	LMMM	STBB	SBBB	STBM	SBBM	STMB	SBMB	STMM	SBMM	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
stdev	33.29	16.00	13.64	1.86	10.17	0	0	17	40	4	13	79	0	0	37	8	37	0	18	3	18	24	0	0	298
sdres	16.28	0.51	6.67	0.91	4.97	0	0	8	19	2	6	39	0	0	18	4	18	0	9	2	9	12	0	0	146
maxdd	65.49	2.06	26.83	3.66	20.00	0	0	34	78	8	26	155	0	0	73	15	73	0	35	7	35	47	0	0	586

EXPONENTIAL DISTRIBUTION CONFIDENCE LEVEL 16.97% EQUAL TO MAXIMUM DRAWDOWN OF $ 1.7 million

Conf	AROR	Std	StdR	StdD	Mdd%	LTBB	LBBB	LTBM	LBBM	LTMB	LBMB	LMMB	LTMM	LBMM	LMMM	STBB	SBBB	STBM	SBBM	STMB	SBMB	STMM	SBMM	SMMM	Net
						ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
95.40	98.81	3.11	40.38	5.53	30.13	0	0	51	118	12	39	234	0	0	110	22	110	0	53	10	53	71	0	0	883
99.74	50.96	1.60	20.93	2.93	15.50	0	0	26	61	6	20	121	0	0	57	12	57	0	27	5	27	37	0	0	456
99.90	44.10	1.38	18.26	2.47	13.47	0	0	23	53	5	17	104	0	0	49	10	49	0	24	5	24	32	0	0	395

Best case has 11 of 19 assets.
Total run time 85 mins 6 secs.

FIGURE 24.7 *(Continued)*

StdD is the semi-variance, the standard deviation of the daily drawdowns.

Mdd% is the maximum drawdown as a percentage of the investment.

Column 1 shows that the best solution was found in 59 cycles through the program. The missing numbers occur when no improvement was found for a cycle or sequence of cycles. Allocations shift considerably in the first 15 cases, then stabilize and improve gradually. The objective function gained 80% of its improvement in the first 15 cases, and the remaining 20% in the next 45 cases, and nothing afterwards.

The final solution, which appears in case 59, is rescaled at the bottom of part (a) to the targeted confidence level of 99.75 (less than a 0.25% chance of losing the targeted amount of 16.97% of the investment during the test interval of 7.16 years). The annualized rate of return is 49.46% for this solution.

The results of the three passes show that the best objective function increases in each pass from 17.315 to 17.801, and finally to 17.894. The rate of return (seen as *AROR* in column 2 in the last 3 lines of each pass), however, only improves from 49.46% to 50.95%, showing that returns are very similar. Comparing Pass 1 and Pass 3, the allocations might seem very different; however, the same strategies receive the allocations, and the distribution of the allocations is similar. In Pass 1, the strategy LMMB receives 107 contracts (confidence level 99.74), and in Pass 3 it is given 121 contracts. At the same time, strategy SBBB drops from 69 to 57 contracts. There are clearly a range of allocations that will return similar results. In each of the three passes, the number of assets allocated differed from 10 to 12, although all were no higher than the requested level of 12.

The total elapsed time for the test of 3 passes, each with 100 cases (or cycles), was 85 minutes on a 1.7 mHz Pentium 4 processor, the best available in 2004, but slow compared to today's standards. Considering the number of possible combinations, the time is not a deterrent.

Test Interval for Portfolio Allocation

Short test intervals are responsive to current market conditions but do not reflect the full balance of bull and bear markets. Statisticians always prefer more data to less data. Rolling a 6-month portfolio every month limits the view of each performance or market to a very narrow window. No one has yet shown that an allocation based on short intervals is better than one based on a longer period. In order to include a fair sample of bull and bear market periods, it is necessary to begin with data from about 1995, capturing 3 years of bull market performance before the end of the tech bubble, then an incredibly volatile and unusual period through the writing of this edition.

The period over which a portfolio is specified must also include the full seasonal pattern of the assets; otherwise there can be unequal risk in seasonal and nonseasonal assets that are included at the same time. This would be the case with the heating oil and unleaded gasoline futures, or resort stocks where one company benefited from winter

business and the other from summer business. The market that is out of season will show lower risk. The two opposed markets may be shifting risk in a way that is always out of phase with portfolio allocations that are based on the past 6 months of data.

VOLATILITY STABILIZATION

When a portfolio is engineered, it has a specific risk profile. Normally, this can be stated as a 12% volatility, equal to 1 standard deviation of the annualized returns, indicating that there is a 16% chance that the drawdown will be greater than 12%, and an 8% chance that it will be greater than 24%. However, these statistics are calculated over many years of data, and the averages do not explain the variation in risk over that period. It is possible with "12% volatility" that there will be years of 8% volatility and some uncomfortable spikes of 18% or even 25%. Investors would like more stability.

Volatility stabilization has become an important part of the entire managed account venue. There are many ways that it can be accomplished, and all of them will control the fluctuation of risk to some degree. None of the methods eliminate the need for an additional risk control in the form of value-at-risk. Volatility stabilization looks only at historic volatility while VaR tries to anticipate risk.

Capping leverage is another forward-looking risk control. If you prevent high leverage, then you prevent high risk. Some traders and analysts believe that you limit returns too much when leverage is capped, so it is often not used. Risk managers should not discard the idea without thoroughly testing it.

As an example of volatility stabilization, consider a rolling measure of annualized volatility. Table 24.8 is the beginning of a systematic trading performance showing the NAVs and the daily returns in the first 3 columns. This is the output from a portfolio of mixed assets with no regard to the volatility, which is annualized at 8.52% over about 17 years, but the range is from 1.5% to 25.5%, with an annualized return of 10.5%. If an investor wanted more volatility, say 12%, we would expect the returns to be increased proportionally.

To accomplish an increase in volatility, we could simply multiply all the returns by a volatility factor (VF)

$$VF = \frac{Target\ volatility}{Annualized\ volatility}$$

In this case $VF = 0.12/.0832 = 1.442$. Then all positions traded would have been increased by a factor of 1.442 as would the returns. But then the factor used is dependent on the amount of history and the day we perform the calculation.

Instead, it is more practical to calculate a rolling annualized volatility and adjust the position sizes by the lagged volatility factor. In the example, if we use a 10-day rolling annualized volatility, the new NAVs compared to the original are shown in Figure 24.8 and the corresponding rolling volatility in Figure 24.9. The volatility is shown only for

TABLE 24.8 Example of Volatility Stabilization

Date	Original NAVs	Returns	Ann Vol	Vol Factor	Adj Returns	Vol-Adj NAVs	Ann Adj Vol
4/20/1994	100.000			0.12			
4/21/1994	99.645	−0.00355					
4/22/1994	99.667	0.00022					
4/25/1994	99.821	0.00154					
4/26/1994	99.691	−0.00130					
4/27/1994	99.583	−0.00108					
4/28/1994	99.896	0.00314					
4/29/1994	99.897	0.00001					
5/2/1994	100.366	0.00470					
5/3/1994	100.378	0.00011					
5/4/1994	100.320	−0.00057	0.0371	3.235		100	
5/5/1994	100.124	−0.00196	0.0331	3.629	−0.0063	99.366	
5/6/1994	100.862	0.00737	0.0477	2.515	0.0267	102.024	
5/9/1994	101.300	0.00435	0.0503	2.386	0.0109	103.139	
5/10/1994	100.511	−0.00779	0.0678	1.769	−0.0186	101.222	
5/11/1994	100.917	0.00404	0.0687	1.748	0.0071	101.945	
5/12/1994	101.021	0.00103	0.0679	1.767	0.0018	102.129	
5/13/1994	100.988	−0.00033	0.0681	1.762	−0.0006	102.070	
5/16/1994	100.713	−0.00272	0.0673	1.783	−0.0048	101.580	
5/17/1994	100.237	−0.00472	0.0720	1.667	−0.0084	100.724	
5/18/1994	100.533	0.00295	0.0735	1.632	0.0049	101.220	0.1963
5/19/1994	100.389	−0.00143	0.0731	1.641	−0.0023	100.983	0.1930
5/20/1994	100.968	0.00576	0.0691	1.736	0.0095	101.938	0.1432
5/23/1994	101.617	0.00644	0.0734	1.636	0.0112	103.077	0.1437
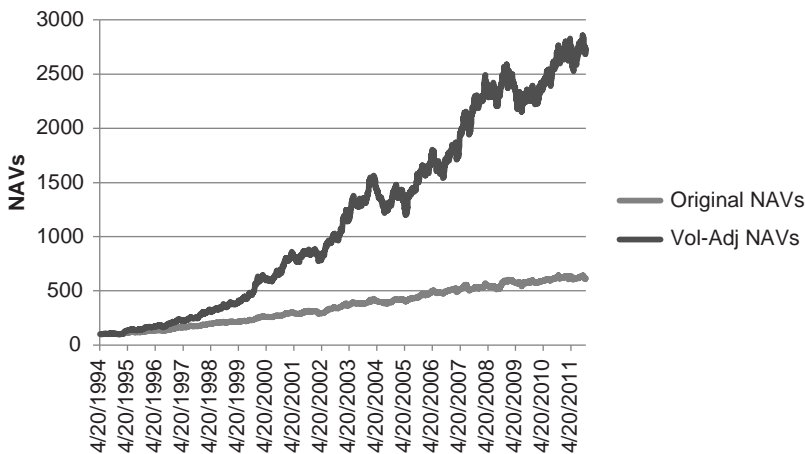
FIGURE 24.8 Original NAVs at an annualized volatility of 8.32% and the volatility-adjusted NAVs at 13.7%.

a short period in the second half of 2011; however, it can be seen as more uniform than the original values. The average of the adjusted volatility is 13.7%, exceeding the target of 12%. This will happen when volatility generally increases over the test period because the volatility factor is a lagged function.

A benefit of volatility adjusting, besides the need to achieve a more predictable leverage level, is the improvement in overall performance. The original NAVs had an information ratio of 1.27 over 17 years, while the adjusted NAVs had a ratio of 1.38.

Switching Costs

Each time the volatility factor changes, the position size being traded must be adjusted. This can incur significant costs. The way to deal with this is to:

- Use a longer calculation period for the volatility calculation.
- Use a volatility change threshold, so that no positions are altered unless the volatility changes by, for example, 20%.

Deciding on the right combination is a matter of testing a specific strategy.

Capping Exposure

Registered funds, in particular the European UCITS requirements, can restrict leverage by imposing restrictions on exposure. This mainly affects managed futures. The requirements are:

- Total exposure cannot exceed 2.5 × current account equity.
- Exposure to commodities cannot exceed 20% of current account equity.
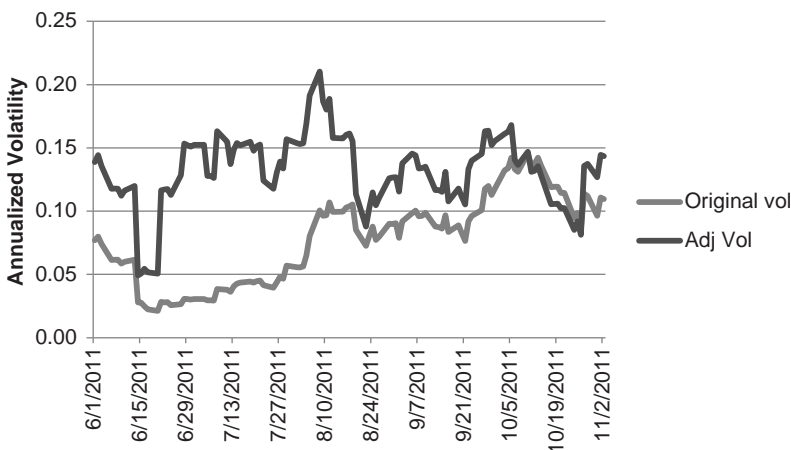
FIGURE 24.9 Rolling annualized volatility of the original and the adjusted NAVs.

The severe restrictions for commodity trading can be explained by fear of delivery, even though some financial instruments have physical delivery.

Capping must be applied after all volatility stabilization and other adjustment factors have been accounted for. It is the last step. To understand the effect of capping, we must first explain exposure. *Exposure* is the face value of the assets being held. For equities, it is simply the number of shares times the share price. For commodities, it is the number of contracts times the *big point value*. This is the same as the face value except for interest rates with maturities less than one year. For example, with 3-month maturities the face value of $1,000,000 is divided by 4.

Capping must be concurrent with taking positions. In a manner similar to controlling VaR, you must evaluate the exposure relative to the cap before setting positions that would exceed the cap. If the exposure E_t is greater than the cap, C, then the normal process is to reduce all positions by the ratio C/E_t, including the new positions that caused the exposure to exceed the cap. By reducing all positions in the portfolio, you maintain the portfolio balance and its integrity.

Capping should reduce the total returns of the portfolio but it may have only minimal effect on the information ratio. High leverage is usually associated with low volatility, and low volatility is associated with lack of price direction, so there is only a small expectation of reduced performance. Given regulations, most portfolio managers have no choice.

Statistical Tables

PROBABILITY DISTRIBUTIONS TABLES

TABLE A1.1   Normal Curve Areas
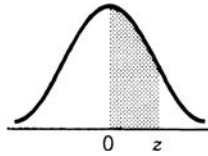
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3820
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

TABLE A1.2 T-Distribution

df \ α	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

Source: Paul G. Hoel, *Elementary Statistics,* 3rd ed. (New York: John Wiley & Sons, 1971).

TABLE A1.3 Uniform Random Numbers

	1	2	3	4	5	6	7	8	9	10	11
1	10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969
2	22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666
3	24130	48360	22527	97265	76393	64809	15179	24830	49340	32081	30680
4	42167	93093	06243	61680	07856	16376	39440	53537	71341	57004	00849
5	37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110
6	77921	06907	11008	42751	27756	53498	18602	70659	90655	15033	21916
7	99562	72905	56420	69994	98872	31016	71194	18738	44013	44840	63213
8	96301	91977	05463	07972	18876	20922	94595	56869	69014	60045	18425
9	89579	14342	63661	10281	17453	18103	57740	84378	25331	12566	58678
10	85475	36857	53342	53988	53060	59533	38867	62300	08158	17893	16439
11	28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	91547
12	63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234
13	09429	93969	52636	92737	88974	33488	36320	17617	30015	07272	84115
14	10365	61129	87529	85689	48237	52267	67689	93394	01511	26358	85104
15	07119	97336	71048	08178	77233	13916	47564	81056	97735	85977	29372
16	51085	12765	51821	51259	77452	16308	60756	92144	49442	53900	70960
17	02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64835
18	01011	54092	33362	94904	31273	04146	18594	29852	71585	85030	51132
19	52162	53916	46369	58586	23216	14513	83149	98736	23495	64350	94738
20	07056	97628	33787	09998	42698	06691	76988	13602	51851	46104	88916
21	48663	91245	85828	14346	09172	30168	90229	04734	59193	22178	30421
22	54164	58492	22421	74103	47070	25306	76468	26384	58151	06646	21524
23	32639	32363	05597	24200	13363	38005	94342	28728	35806	06912	17012
24	29334	27001	87637	87308	58731	00256	45834	15398	46557	41135	10367
25	02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586
26	81525	72295	04839	96423	25878	82651	66566	14778	76797	14780	13300
27	29676	20591	68086	26432	46901	20849	80768	81536	86645	12659	92259
28	00742	57392	39064	66432	84673	40027	32832	61362	98947	96067	64760
29	05366	04213	25669	26422	44407	44048	37937	63904	45766	66134	75470
30	91921	26418	64117	94305	26766	25940	39972	22209	71500	64568	91402
31	00582	04711	87917	77341	42206	35126	74087	99547	81817	42607	43808
32	00725	69884	62797	56170	86324	88072	76222	36086	84637	93161	76038
33	69011	65795	95876	55293	18988	27354	26575	08625	40801	59920	29841
34	25976	57948	29888	88604	67917	48708	18912	82271	65424	69774	33611
35	09763	83473	73577	12908	30883	10317	28290	35797	05998	41688	34952
36	91567	42595	27958	30134	04024	86385	29880	99730	55536	84855	29080
37	17955	56349	90999	49127	20044	59931	06115	20542	18059	02008	73708
38	46503	18584	18845	49618	02304	51038	20655	58727	28168	15475	56942
39	92157	89634	94824	78171	84610	82834	09922	25417	44137	48413	25555
40	14577	62765	35605	81263	39667	47358	56873	56307	61607	49518	89565

Source: Francis F. Martin, *Computer Modeling and Simulation* (New York: John Wiley & Sons, 1968), 288. Reprinted with permission of John Wiley & Sons, Inc.

Matrix Solution to Linear Equations and Markov Chains

DIRECT SOLUTION AND CONVERGENCE METHOD

Before computer programs offered ready solutions, problems such as Markov chains were solved in a direct manner, by algebraically manipulating the equations. This direct solution requires an understanding of simple matrix arithmetic and very careful attention to calculating the numbers correctly. The convergence method is now easier, although it requires many more calculations. Without the computer, we would never even consider using this approach—with a computer, it is the best choice.

GENERAL MATRIX FORM

A matrix is a rectangular arrangement of elements into rows and columns. Matrix A is said to be $m \times n$ (pronounced "m by n") if there are m rows and n columns in A.

$$a_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Certain properties of a matrix make it a valuable tool for solving simultaneous linear equations. These elementary matrix operations, called *transformations*, allow you to alter the rows (which will represent equations) without changing the solution. There are three basic row operations:

1. Multiplication or division of all elements of the row by any number.

2. Interchanging of any two rows (and consequently of all rows).

3. The addition or subtraction of the elements of one row with the corresponding elements of another.

To relate the matrix to simultaneous linear equations, consider a three-equation example, where only the coefficients, x, remain as unknowns:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = a_{14} \qquad (1)$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = a_{24} \qquad (2)$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = a_{34} \qquad (3)$$

The three elementary row operations can now be interpreted in terms of simple operations on an equation:

1. When both sides of an equation are multiplied or divided by the same value, the results are equal.
2. Any two equations in a system of equations can be interchanged with no effect.
3. When two equals are added or subtracted, the results are equal.

DIRECT SOLUTION

Putting these rules into use, write the coefficients of the three simultaneous linear equations as a 3×4 coefficient matrix,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$$

with the objective of reducing the matrix to the form

$$\begin{pmatrix} 1 & 0 & 0 & A_1 \\ 0 & 1 & 0 & A_2 \\ 0 & 0 & 1 & A_3 \end{pmatrix}$$

which would mean that $x_1 = A_1$, $x_2 = A_2$, and $x_3 = A_3$ since

$$1 \times x_1 + 0 \times x_2 + 0 \times x_3 = A_1$$
$$0 \times x_1 + 1 \times x_2 + 0 \times x_3 = A_2$$
$$0 \times x_1 + 0 \times x_2 + 1 \times x_3 = A_3$$

To achieve the results, divide the first equation by a_{11} the first element, leaving

$$\begin{pmatrix} 1 & \dfrac{a_{12}}{a_{11}} & \dfrac{a_{13}}{a_{11}} & \dfrac{a_{14}}{a_{11}} \end{pmatrix} \qquad (2\text{-}1)$$

and then multiply by a_{21} to get the first elements in rows one and two the same:

$$\left(a_{21} \quad \frac{a_{12} \times a_{21}}{a_{11}} \quad \frac{a_{13} \times a_{21}}{a_{11}} \quad \frac{a_{14} \times a_{21}}{a_{11}} \right) \qquad (2\text{-}2)$$

Now subtract (1-2) from (2) and get

$$\left(0 \quad a_{22} - \frac{a_{12} \times a_{21}}{a_{11}} \quad a_{23} - \frac{a_{13} \times a_{21}}{a_{11}} \quad a_{24} - \frac{a_{14} \times a_{21}}{a_{11}} \right) \qquad (2\text{-}3)$$

That successfully eliminates the first element a_{21} from the second equation. By going back and multiplying (2-1) by a_{31} and subtracting the resulting equation from (3), a_{31} can be eliminated from equation (3). Now column 1 looks like

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Column two or any column can be operated upon in the same manner as the first column:

1. Divide row n by the element in position n (for example, a_{nn}), thereby setting $a_{nn} = 1$.
2. Multiply row n by the corresponding element in row $i \neq n$, so that $a_{nn} \times a_{in} = a_{in}$.
3. Subtract row n from row i, resulting in $a_{in} = 0$, and all other elements reduced by the corresponding element in row n.

Continue this procedure for each row i until all elements

$$a_{1n}, a_{2n}, \ldots, a_{i-1n}, a_{1+1n}, \ldots, a_{mn} = 0 \quad \text{and} \quad a_{in} = 1$$

For example, beginning with

$$\begin{pmatrix} 2 & 4 & 8 & 34 \\ 6 & 5 & 2 & 22 \\ 3 & 6 & 5 & 30 \end{pmatrix}$$

Divide row 1 by the value 2 (position a_{11}):

$$\begin{pmatrix} 1 & 2 & 4 & 17 \\ 6 & 5 & 2 & 22 \\ 3 & 6 & 5 & 30 \end{pmatrix}$$

Multiply row 1 by the value 6 to get (6 12 24 102) and subtract the new calculated row from row 2:

$$\begin{pmatrix} 1 & 2 & 4 & 17 \\ 0 & -7 & -22 & -80 \\ 3 & 6 & 5 & 30 \end{pmatrix}$$

Multiply row 1 by the number 3 to get (3 6 12 51), then subtract the calculated row from row 3:

$$\begin{pmatrix} 1 & 2 & 4 & 17 \\ 0 & -7 & -22 & -80 \\ 0 & 0 & -7 & -21 \end{pmatrix}$$

Column 1 is now completed. Divide row 2 by −7 and get:

$$\begin{pmatrix} 1 & 2 & 4 & 17 \\ 0 & 1 & +22/7 & +80/7 \\ 0 & 0 & -7 & -21 \end{pmatrix}$$

Multiply row 2 by 2 and subtract the result (0 2 44/7 160/7) from row 1 to eliminate position 2 in the first row:

$$\begin{pmatrix} 1 & 0 & -16/7 & -41/7 \\ 0 & 1 & +22/7 & +80/7 \\ 0 & 0 & -7 & -21 \end{pmatrix}$$

Because the element in row 3 and column 2 is already 0, move to row 3. Divide row 3 by −7 and get

$$\begin{pmatrix} 1 & 0 & -16/7 & -41/7 \\ 0 & 1 & +22/7 & +80/7 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Multiply row 3 by −16/7 and subtract the result (0 0−16/7 −48/7) from row 1:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & +22/7 & +80/7 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Multiply row 3 by +22/7 and subtract (0 0+22/7 +66/7) from row 2:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

The results show that $x_1 = 1$, $x_2 = 2$, $x_3 = 3$.

There are other ways to reduce the matrix to the final representation, but this technique is well defined and lends itself to being programmed on a computer.

Solution to Weather Probabilities Expressed as a Markov Chain

Let

$$A = P(\text{clear})_{i+1} = P(\text{clear})_i$$
$$B = P(\text{cloudy})_{i+1} = P(\text{cloudy})_i$$
$$C = P(\text{rainy})_{i+1} = P(\text{rainy})_i$$

because when they converge all ith elements will equal $(i+1)$th elements. The equations are

$$A = 0.7A + 0.2B + 0.2C \qquad (1)$$
$$B = 0.25A + 0.6B + 0.4C \qquad (2)$$
$$C = 0.05A + 0.2B + 0.4C \qquad (3)$$

In addition,

$$A + B + C = 1 \qquad (4)$$

To solve this system of equations using matrices, convert and add equation (3) to (4),

$$-0.3A + 0.2B + 0.2C = 0 \qquad (1')$$
$$0.25A - 0.4B + 0.4C = 0 \qquad (2')$$
$$1.05A + 1.2B + 0.4C = 1 \qquad (3')$$

which becomes the matrix

$$\begin{pmatrix} -0.3 & 0.2 & 0.2 & 0 \\ 0.25 & -0.4 & 0.4 & 0 \\ 1.05 & 1.2 & 0.4 & 1 \end{pmatrix} \quad \begin{matrix} (1') \\ (2') \\ (3') \end{matrix}$$

The following are key steps in the solution:

1. Reduce the first row,

$$\begin{pmatrix} 1 & -0.6667 & -0.6667 & 0 \\ 0.25 & -0.4 & 0.4 & 0 \\ 1.05 & 1.2 & 0.4 & 1 \end{pmatrix}$$

and make the leading entries of rows 2 and 3 zero:

$$\begin{pmatrix} 1 & -0.6667 & -0.6667 & 0 \\ 0 & -0.2333 & 0.5667 & 0 \\ 0 & 1.9000 & 1.1000 & 1 \end{pmatrix}$$

2. Reduce the second row,

$$\begin{pmatrix} 1 & -0.6667 & -0.6667 & 0 \\ 0 & 1 & -2.4291 & 0 \\ 0 & 1.9000 & 1.1000 & 1 \end{pmatrix}$$

and make the second entries of rows 1 and 3 zero:

$$\begin{pmatrix} 1 & 0 & 2.2857 & 0 \\ 0 & 1 & -2.4291 & 0 \\ 0 & 0 & 5.7143 & 1 \end{pmatrix}$$

3. Reduce the third row,

$$\begin{pmatrix} 1 & 0 & 2.2862 & 0 \\ 0 & 1 & -2.4291 & 0 \\ 0 & 0 & 1 & 0.1750 \end{pmatrix}$$

and make the third entry of rows 1 and 2 zero:

$$\begin{pmatrix} 1 & 0 & 0 & 0.4000 \\ 0 & 1 & 0 & 0.4250 \\ 0 & 0 & 1 & 0.1750 \end{pmatrix}$$

Then $A = 0.4000$, $B = 0.4250$, and $C = 0.1750$.

Computer Program Direct Solution

The following FORTRAN program accepts a system of 10 equations in 10 variables (unknowns) and applies the matrix method of solution. This is done using the method of Gaussian elimination.

```
      PROGRAM MATRIX
C----  Matrix solution to simultaneous linear equations
C----  Copyright 1986 PJ Kaufman
      DIMENSION A(10,10), C(10)
      OPEN(6,FILE='PRN')
      WRITE(*,7000)
 7000 FORMAT(' Enter matrix size (n)>'\)
      READ(*,5000)N
 5000 FORMAT(BN,14)
      IF(N.GT.1O)STOP 'Matrix limited to 10 x 10¢
      WRITE(*,7001) (1,1=1,N)
```

```
7001   FORMAT(' Enter matrix elements row by row under headings
      +    ' do not include constant to right of
      +    7X,10(12,'-xxxxx'))
       DO 20 I = 1,N
         WRITE(*,7003)1
7003     FORMAT(' Row',12,':'\)
         READ(*,5003) (A(IMP), J=1,N)
5003     FORMAT(BN,10F8.0)
  20     CONTINUE
       WRITE(*,7004) (J,J=1,N)
       WRITE(6,7004) (J,J=1,N)
7004   F0RMAT(/' Input matrix is:'//9X,10(I2,'-col '))
       DO 30 I = 1,N
          WRITE(6,7005)I,(A(I,J),J=1,N)
  30      WRITE(*,7005)I,(A(I,J),J=1,N)
7005     FORMAT(/' Row',I2,':',10F8.3)
       WRITE(*,7008)(I,I=1,N)
7008   FORMAT(/' Enter constant vector under headings...'/
      +    1X,10(I1,'xxxxxx '))
         READ(*,5008)(C(I),I=1,N)
5008   FORMAT(BN,10F8.0)
       WRITE(*,7009)(C(I),I=1,N)
       WRITE(6,7009)(C(I),I=1,N)
7009   FORMAT(/' Constant vector is:'/10F8.3)
C----- Process row by row (Gaussian Elimination)
       DO 100 I = 1,N
          DIV = A(I,I)
          DO 40 J = 1,N
  40       A(I,J) = A(I,J)/DIV
          C(I) = C(I)/DIV
C----- Zero out column I for each row
          DO 60 J = 1,N
             IF(J.EQ.I)GOTO 60
             FACT0R = A(J,I)
             DO 50 K = 1,N
  50            A(J,K) = A(J,K) - A(I,K)*FACT0R
             C(J) = C(J) - C(I)*FACT0R
  60         CONTINUE
 100     CONTINUE
       WRITE(*,7007)(C(I),I = 1,N)
       WRITE(6,7007)(C(I),I = 1,N)
7007   FORMAT(/' Solution vector is: '/10F8.3)
       CALL EXIT
       END
```

Sample Computer Printout

```
MATRIX
Enter matrix size (N)> 3
Enter matrix elements row by row under headings...
do not include constant to right of =
      1-xxxxx 2-xxxxx 3-xxxxx
Row 1: 2       4       8
Row 2: 6       5       2
Row 3: 3       6       5
Input matrix is:
      1-col    2-col    3-col
Row 1: 2.000    4.000    8.000
Row 2: 6.000    5.000    2.000
Row 3: 3.000    6.000    5.000
Enter constant vector under headings.,
1xxxxxx 2xxxxxx 3xxxxxx
34      22       30
Constant vector is:
  34.000    22.000    30.000
Solution vector is:
   1.000     2.000     3.000
```

CONVERGENCE METHOD

This method performs a series of matrix multiplications until the difference between the new and previous matrix is very small.[1] Following the procedure in Chapter 2, Basic Concepts and Calculations, we can create a 3×3 frequency matrix by counting the number of up, down, and neutral days that follow other up, down, and neutral days. We use the term *neutral* to allow very small price changes to be considered in this group, rather than limit it to only those days with zero changes. For this example we will look at the number of up, down, and neutral days that follow a 5-day trend that was considered up, down, or neutral on the previous day. Suppose the results were those shown in the *frequency matrix F*:

		Next Day Price Change			
		Up	**Neutral**	**Down**	**Total**
Previous day trend	Up	70	40	30	140
	Neutral	50	30	45	125
	Down	35	45	65	145

---

[1] George R. Arrington, "Markov Chains," *Technical Analysis of Stocks & Commodities* (December 1993).

Divide each item in *F* by the total given at the end of that row, and get the probability of each occurrence in a *transition matrix T*.

		Next Day Price Change		
		Up	**Neutral**	**Down**
Previous day trend	Up	0.500	0.285	0.214
	Neutral	0.400	0.240	0.360
	Down	0.241	0.310	0.448

Now it is necessary to perform matrix multiplication. To multiply matrix *A* by matrix *B*, we multiply the corresponding items in row *i* of *A* by the corresponding item in column *j* of *B*, add those products together to get the item in row *i*, column *j* of the new matrix *C*. If we have two 3×3 matrices *A* and *B*, and we wanted to find the element in row 2, column 1 of the new product matrix *C*, we would multiply and add

$$c_{21} = a_{21} \times b_{11} + a_{22} \times b_{21} + a_{23} \times b_{31}$$

The general formula for this is

$$c_{ij} = \sum_{k=1}^{N} (a_{ik} \times b_{kj})$$

For spreadsheet users, there is no general formula to copy from one cell to another. For matrix *A* of 3 columns and 3 rows located in columns *A*, *B*, *C*, rows 1, 2, 3; matrix *B* in columns *D*, *E*, *F*, rows 1, 2, 3; and matrix *C* in rows *G*, *H*, *I*, columns 1, 2, 3, we enter the formula for c_{11} in cell *G*1 as

$$\textit{cell } G1 = A1*D1 + B1*D2 + C1*D3$$

While this is clearly tedious, at least the arithmetic will be correct.

Iterative Matrix

We can now find the solution to the Markov chain, the long-term probabilities, by performing a series of matrix multiplications beginning with the transition matrix.

1. Multiply the transition matrix *T* by itself to get the first iterative matrix I_1,

$$I_1 = T \times T$$

2. Multiply the iterative matrix by the transition matrix to get the next iterative matrix,

$$I_2 = I_1 \times T$$

3. Continue to multiply the last iterative matrix by the transition matrix until the new iterative matrix is unchanged (or very close) to the previous iterative matrix,

$$I_n = I_{n-1} \times T$$

When each element satisfies the condition

$$\text{abs}(I_{ij}(n) - I_{ij}(n-1)) < 0.001$$

then the iterative matrix holds the final long-term probabilities of the Markov chain.

Trigonometric Regression for Finding Cycles

The following computer programs and examples solve the cycle problems in Chapter 11.

SINGLE-FREQUENCY TRIGONOMETRIC REGRESSION

The FORTRAN program *TRIG1* and the subroutine *LINREG* are used to find the single-frequency representation of the copper cycle. The program output is clearly separated into the following information:

1. Input data, where each time period is the average cash price for a calendar quarter.
2. The solution to the linear regression, giving the detrending line. With $b = 0.267$, there is an inflationary bias of $+0.267¢$ per quarter.
3. The detrended data resulting from subtracting the line values (2) from the original data (1).
4. Intermediate values for α, ω, and T.
5. The constant values a and b for the normal equations solving the single-frequency problem.
6. The cycle resulting from the detrended data.
7. The final cycle with the trend added back.

The results show a copper cycle of approximately 8.4 quarters, or slightly more than 2½ years.

An additional test was run on monthly cash corn prices from 1964 through 1983 to see if the seasonal cycle dominated the detrended pattern. The linear regression equation used for detrending was calculated as:

$$y = 0.939 + .01x$$

APPENDIX 3

showing only a 1¢ per bushel per month rate of inflation, despite the bull markets in 1973 and 1980 through 1981. The cycle showed a period of 21.4 months, with the last highs in the cycle in August 1983 and the last lows in September 1982. Because this is clearly not a seasonal pattern, it must be either:

1. Dominated by other supply-demand characteristics, such as stocks, or,
2. Distorted by the nonseasonal rallies of 1973 and 1980, which each took three years to return to the traditional seasonal pattern.

```
      PROGRAM TRIG1
C---- Single Frequency Trigonometric Regression
C---- Copyright 1986 PJ Kaufman
      DIMENSION X(250), Y(250), D(250), R(250)
      DOUBLE PRECISION SC2,SCD
      DATA MAX/25O/
      OPEN(6,FILE='PRN')
C---- Read data, set X to incremental time
C---- I = 1
C---- WRITE(*,7000)
 7000 FORMAT(' SINGLE FREQUENCY TRIGONOMETRIC REGRESSION'/
     +        'Enter data 1 per line'/' Extra <return> ends input'//
     +        'Enter below:')

   10 WRITE(*,7001)I
 7001 FORMAT(14,'>'\)
      READ(*,5000)Y(I)
 5000 FORMAT(BN,F8.2)
      X(I) = I
      IF(Y(I).NE.0)THEN
        IF(I.GT.MAX)STOP 'Maximum data exceeded'
        I = I+1
        GOTO 10
        ENDIF
      N = I-1
      WRITE(6,6000)
 6000 FORMAT('1Single Frequency Trigonometric Regression',2OX,
     +        'Data Input'//' Time',16X,'Prices')
      DO 40 J = 1,N,4
   40    WRITE(6,6001)J,(Y(I),I=J,J+3)
 6001    FORMAT(14,4F8.2)
C---- Linear regression analysis for detrending
      CALL LINREG(N,Y,A,B,SD)
      WRITE(6,6002)A,B
 6002 FORMAT(/'Linear regression results: A =',F8.3,', B =',F8.3)
```

```
C----  Detrend data into D and computer sums for equation (4)
       DO 60 I = 1,N
   60     D(I) = Y(I) - (A+B*I)
C----  Print detrended data
       WRITE(6,6003)
 6003  FORMAT(/' Detrended data'/)
       DO 65 I = 1,N,4
   65     WRITE(6,6001)I,(D(J),J=I,I+3)
       SC2 = 0
       SCD = 0
C----  Solve equation (4) using detrended data
       DO 70 I = 2,N-1
          DI = D(I)
          SC2 = SC2 + DI*DI
   70     SCD = SCD + DI*(D(I-1)+D(I+1))

       ALPHA = SCD/SC2
       WRITE(6,6004)SC2,SCD,ALPHA
 6004  FORMAT(/'Sum C-squared =',F8.1,', Sum C x D =',F8.1,', Alpha =',
      +     F8.3)
C----  Solve for omega
       W = ACOS(ALPHA/2)
       T = 36O/W
       WRITE(6,6O14)W,T
 6014  FORMAT(/'Omega (W) =',F5.1,' degrees, Period (T) =',F6.2,
      +       'time units')

C----  Sums for normal equations
       COS2 = 0
       COSSIN = 0
       YCOS = 0
       SINCOS = 0
       SIN2 = 0
       YSIN = 0
       DO 8O I = 1,N
          C = COS(W*I)
          S = SIN(W*I)
          COS2 = COS2 + C*C
          COSSIN = COSSIN + C*S
          YCOS = YCOS + Y(I)*C
          SINCOS = SINCOS + S*C
          SIN2 = SIN2 + S*S
   80  YSIN = YSIN + Y(I)*S
C----  Solve normal equations
       TB = (YSIN*COS2 - YCOS)/(SIN2*COS2 - COSSIN)
       TA = (YCOS - B*COSSIN)/COS2
       WRITE(6,6005)TA,TB
```

```
 6005  FORMAT(/' Solution to normal equations: A =',F8.3,', B =',F8.3)
C----  Values of fitted curve using detrended data
       DO 90 I = 1,N
   90     R(I) = TA*COS(W*I) + TB*SIN(W*I)
       WRITE(6,6006)
 6006  FORMAT(/' Trigonometric regression results using detrended data',
      +       /)
       DO 100 I = 1,N,4
  100    WRITE(6,6001)I,(R(J),J=I,I+3)
C----  Add trend back to result
       DO 110 I = 1,N
  110    R(I) = R(I) + A + B*I

       WRITE(6,6007)
 6007  FORMAT(/' Final regression results with trend added back'/)
       DO 120 I = 1,N,4
  120    WRITE(6,6001)I, (R(J),J=I,I+3)

       CALL EXIT
       END

       SUBROUTINE LINREG(N,DATA,A,B,SD)
C----  Generalized simple linear regression
       DIMENSION DATA(2)

C----  Initialize sums
       SX=0.
       SY=0.
       SXY=0.
       SX2=0.
       A=0.
       B=0.
       SD=0.
       IF(N.LE.2)RETURN

       DO 100 I=2,N
          X = I
          Y = DATA(I)
          SX=SX+X
          SY=SY+Y
          SXY=SXY+Y*X
          SX2=SX2+X*X
  100    CONTINUE
       M=N-1
       B=(M*SXY-SX*SY) / (M*SX2-SX*SX)
       A=(SY-B*SX) / M
```

```
C----- Residuals
      SSR=0
      DO 200 I=2,N
         Y=DATA(I)
         R=Y-(A+B*I)
         SSR=SSR+R*R
  200   CONTINUE
      SD=SQRT(SSR/M)
      RETURN
      END
```

Single Frequency Trigonometric Regression

Time		Prices		
1	22.12	22.46	22.17	22.00
5	23.18	24.56	25.57	30.59
9	28.23	33.77	35.90	40.05
13	46.22	51.48	40.76	40.16
17	36.51	29.30	30.36	36.42
21	39.75	30.07	29.08	32.13
25	38.94	42.95	43.38	46.23
29	47.70	46.98	35.78	27.35
33	25.40	29.45	27.15	28.48
37	32.74	33.53	30.01	29.25
41	36.82	45.07	55.13	65.51
45	66.56	70.06	27.30	35.62
49	32.06	31.46	35.75	36.46
53	38.22	43.24	45.46	38.96
57	37.08	38.72	34.01	33.00
61	35.07	40.23	41.63	44.95
65	51.12	63.71	59.56	63.38

Linear regression results: A = 28.889, B = .267

Detrended data

1	-7.04	-6.96	-7.52	-7.96
5	-7.05	-5.93	-5.19	-.44
9	-3.06	2.21	4.07	7.95
13	13.86	18.85	7.86	7.00
17	3.08	-4.40	-3.61	2.19
21	5.25	-4.70	-5.96	-3.17
25	3.37	7.11	7.28	9.86
29	11.06	10.07	-1.39	-10.09
33	-12.31	-8.52	-11.09	-10.03
37	-6.04	-5.51	-9.30	-10.33
41	-3.02	4.96	14.75	24.86
45	25.65	28.88	-14.15	-6.10
49	-9.92	-10.79	-6.77	-6.32
53	-4.83	-.08	1.87	-4.89
57	-7.04	-5.67	-10.64	-11.92
61	-10.12	-5.23	-4.09	-1.04
65	4.86	17.19	12.77	16.32

```
Sum C-squared = 6338.4, Sum C x D = 9282.2, Alpha = 1.464
Omega (W) =  .7 degrees, Period (T) =480.50 time units
Solution to  normal equations: A = -.603, B = 1.831
Trigonometric regression results using detrended data
```

1	.81	1.78	1.80	.86
5	-.54	-1.66	-1.88	-1.10
9	.27	1.50	1.92	1.32
13	.01	-1.31	-1.92	-1.51
17	-.28	1.09	1.88	1.66
21	.56	-.85	-1.80	-1.79
25	-.82	.59	1.68	1.87
29	1.06	-.32	-1.53	-1.92
33	-1.28	.04	1.34	1.92
37	1.47	.23	-1.13	-1.89
41	-1.64	-.51	.89	1.82
45	1.77	.77	-.64	-1.71
49	-1.86	-1.02	.37	1.56
53	1.91	1.24	-.09	-1.38
57	-1.93	-1.44	-.19	1.17
61	1.90	1.61	.46	-.94
65	-1.83	-1.75	-.72	.69

```
Final regression results with trend added back
```

1	29.96	31.21	31.50	30.82
5	29.68	28.83	28.88	29.93
9	31.57	33.06	33.75	33.41
13	32.37	31.32	30.98	31.66
17	33.15	34.79	35.85	35.90
21	35.06	33.92	33.23	33.52
25	34.75	36.43	37.79	38.24
29	37.70	36.58	35.64	35.52
33	36.43	38.02	39.59	40.43
37	40.25	39.28	38.18	37.69
41	38.21	39.60	41.27	42.46
45	42.68	41.95	40.81	40.01
49	40.12	41.23	42.89	44.34
53	44.96	44.56	43.49	42.47
57	42.19	42.95	44.47	46.09
61	47.09	47.07	46.18	45.05
65	44.42	44.78	46.07	47.74

TWO-FREQUENCY TRIGONOMETRIC REGRESSION

The FORTRAN program *TRIG2* and its subroutines *LINREG* (found in the previous section) and *MTX* are used to find the two-frequency representation of the copper cycle. The program output is clearly separated into the following steps.

1. Input data, where each time period is the average cash price for a calendar quarter.

2. The solution to the linear regression, giving the detrending line.

3. The detrended data resulting from subtracting the line values (2) from the original data (1).

4. Intermediate values for α_1, α_2, ω_1, and ω_2.

5. Resulting values a_1, b_1, a_2, and b_2, which are derived from the matrix solution using Gaussian elimination.

6. The cycle resulting from the detrended data.

7. The final cycle with the trend added back.

```
      PROGRAM TRIG2
C----  2-Frequency Trigonometric Regression
C----  Copyright 1986 PJ Kaufman
      DIMENSION X(250), Y(250), D(250), R(250), A(4,4), B(4)
      EQUIVALENCE (A(1,1),C2W1),(A(1,2),CW1SW1),(A(1,3),CW1CW2),
     +            (A(1,4),CW1SW2),(A(2,1),DUP1),(A(2,2),S2W1),
     +            (A(2,3),SW1CW2),(A(2,4),SW1SW2),(A(3,1),DUP2),
     +            (A(3,2),DUP3),(A(3,3),C2W2),(A(3,4),CW2SW2),
     +            (A(4,1),DUP4),(A(4,2),DUP5),(A(4,3),DUP6),
     +            (A(4,4),S2W2),(B(1),YCW1),(B(2),YSW1),(B(3),YCW2),
     +            (B(4),YSW2),(B(1),A1),(B(2),B1),(B(3),A2),(B(4),B2)
      DATA MAX/250/,NDIM/4/
      OPEN(6,FILE='PRN')

C----  Read data, set X to incremental time
      I = 1
      WRITE(*,7000)
 7000 FORMAT(' 2-FREQUENCY TRIGONOMETRIC REGRESSION'/
     +      ' Enter data 1 per line'/'Extra <return> end input'//
     +      ' Enter below:')

   10 WRITE(*,7001)I
 7001 FORMAT(14,'>'\)
      READ(*,5000)Y(I)
 5000 FORMAT(BN,F8.2)
      X(I) = I
      IF(Y(I).NE.0)THEN
          IF(I.GT.MAX)STOP 'Maximum data exceeded'
```

```
          I = I + 1
          GOTO 10
          ENDIF
      N = I-1
      WRITE(6,6000)
 6000 FORMAT('12-Frequency Trigonometric Regression'//
      +      'Time',16X,'Prices')
      DO 40 J = 1,N,4
   40   WRITE(6,6001)J,(Y(I),I=J,J+3)
 6001   FORMAT(14,4F8.2)

C---- Linear regression analysis for detrending
      CALL LINREG(N,Y,ALIN,BLIN,SD)
      WRITE(6,6002)ALIN,BLIN

 6002 FORMAT(/' Linear regression results: A =',F8.3,', B =',F8.3)

C---- Detrend data into D and computer sums for equation (4)
      DO 60 I = 1,N
   60   D(I) = Y(I) - (ALIN+BLIN*I)
C---- Print detrended data
      WRITE(6,6003)
 6003 FORMAT(/' Detrended data'/)
      DO 65 I = 1,N,4
   65     WRITE(6,6001)I,(D(J),J=I,I+3)
      SC2 = 0
      SCD = 0
      SCP = 0
      SD2 = 0
      SDP = 0
C---- Solve for alpha1 and alpha2 using detrended data
      DO 70 I = 2,N-3
        C = D(I) + D(I+2)
        T = D(I+1)
        P = D(I-1) + D(I+3)
        SC2 = SC2 + C*C
        SCD = SCD + C*T
        SCP = SCP + C*P
        SD2 = SD2 + T*T
   70   SDP = SDP + T*P
      ALPHA2 = (SDP*SC2-SCP)/(SD2*SC2-SCD)
      ALPHA1 = (SCP-ALPHA2*SCD)/SC2
      T = SQRT(ALPHA1*ALPHA1+8*(1+ALPHA2/2))
      W1 = ACOS((ALPHA1+T)/4)
      W2 = ACOS((ALPHA1-T)/4)
      WRITE(6,6004)SC2,SCD,SCP,SD2,SDP,ALPHA1,ALPHA2,W1,W2
```

```
 6004  FORMAT(/' Intermediate values:'//
      +       ' SUMS C2 =', F8.1,', C*D =',F8.1,', C*P =',F8.1/
      +       'D2 =', F8.1,', D*P =',F8.1//' Alpha1 = ',F8.3,
      +       'Alpha2 =', F8.3,' Omega1 =',F5.2,', Omega2 =',F5.2)
C----  Sums for normal equations. . .to be used for matrix solution
       C2W1 = 0
       CW1SW1 = 0
       CW1CW2 = 0
       CW1SW2 = 0
       YCW1 = 0
       S2W1 = 0
       SW1CW2 = 0
       SW1SW2 = 0
       YSW1 = 0
       C2W2 = 0
       CW2SW2 = 0
       YCW2 = 0
       S2W2 = 0
       YSW2 = 0

       DO 100 I = 1,N
          DI = D(I)
          SW1 = SIN(W1*I)
          CW1 = COS(W1*I)
          SW2 = SIN(W2*I)
          CW2 = COS(W2*I)
          C2W1 = C2W1 + CW1*SW1
          CW1SW1 = CW1SW1 + CW1*SW1
          CW1CW2 = CW1CW2 + CW1*CW2
          CW1SW2 = CW1SW2 + CW1*SW2
          YCW1 = YCW1 + DI*CW1
          S2W1 = S2W1 + SW1*SW1
          SW1CW2 = SW1SW2 + SW1*CW2
          SW1SW2 = SW1SW2 + SW1*SW2
          YSW1 = YSW + DI
          C2W2 = C2W2 + CW2*CW2
          CW2SW2 = CW2SW2 + CW2*SW2
          YCW2 = YCW2 + DI*CW2
          S2W2 = S2W2 + SW2*SW2
  100     YSW2 = YSW2 + DI*SW2
C----  Duplicate calculations for matrix
       DUP1 = CW1SW1
       DUP2 = CW1CW2
       DUP3 = SW1CW2
       DUP4 = CW1SW2
       DUP5 = SW1SW2
       DUP6 = CW2SW2
```

```
      WRITE(6,6009)
 6009 FORMAT(/' Coefficient matrix:'/)
      DO 110 I = 1,NDIM
  110   WRITE(6,6010)(A(I,J),J=1,4),B(I)
 6010   FORMAT(5F8.3)
C----  Solve using matrix Gaussian Elimination
      CALL MTX(A,B,NDIM)


C----  Solution vector
      WRITE(6,6011)(B(I),I=1,NDIM)
 6011 FORMAT(/' Solution vector:'/4F8.3)
C----  Values of fitted curve using detrended data
      DO 90 I = 1,N
   90 R(I) = A1*COS(W1*I) + B1*SIN(W1*I) + A2*COS(W2*I) +
      +      B2*SIN(W2*I)
WRITE(6,6006)
 6006 FORMAT(/' Trigonometric regression results using detrended data',
      +       /)
      DO 105 I = 1,N,4
  105   WRITE(6,6001)I,(R(J),J=I,I+3)
C----  Add trend back to result
      DO 115 I = 1,N
  115 R(I) = R(I) + ALIN + BLIN*I
      WRITE(6,6007)
 6007 FORMAT(/' Final results with trend added back'/)
      DO 120 I = 1,N,4
  120   WRITE(6,6001)I,(R(J),J=I,I+3)


      CALL EXIT
      END


      SUBROUTINE MTX(A,C,N)
C----  Matrix solution to simultaneous linear equations
C----  Copyright 1986 PJ Kaufman

      DIMENSION A(4,4), C(4), A1(4,4), C1(4)

C----  Process row by row (Gaussian Elimination)
      DO 100 I = 1,N
         DIV = A(I,I)

         DO 40 J = 1,N
   40       A(I,J) = A(I,J)/DIV
         C(I) = C(I)/DIV
C----  Zero out column I for each row
         DO 60 J = 1,N
```

```
         IF(J.EQ.I)GOTO 60
         FACTOR = A(J,I)
         DO 50 K = I,N
  50     A(J,K) = A(J,K) - A(I,K)*FACTOR
         C(J) = C(J) - C(I)*FACTOR
  60     CONTINUE

 100     CONTINUE

     RETURN
     END
```

2-Frequency Trigonometric Regression

Time	Prices			
1	22.12	22.46	22.17	22.00
5	23.18	24.56	25.57	30.59
9	28.23	33.77	35.90	40.05
13	46.22	51.48	40.76	40.16
17	36.51	29.30	30.36	36.42
21	39.75	30.07	29.08	32.13
25	38.94	42.95	43.38	46.23
29	47.70	46.98	35.78	27.35
33	25.40	29.45	27.15	28.48
37	32.74	33.53	30.01	29.25
41	36.82	45.07	55.13	65.51
45	66.56	70.06	27.30	35.62
49	32.06	31.46	35.75	36.46
53	38.22	43.24	45.46	38.96
57	37.08	38.72	34.01	33.00
61	35.07	40.23	41.63	44.95
65	51.12	63.71	59.56	63.38

Linear regression results: A = 28.889, B = .267

Detrended data

1	-7.04	-6.96	-7.52	-7.96
5	-7.05	-5.93	-5.19	-.44
9	-3.06	2.21	4.07	7.95
13	13.86	18.85	7.86	7.00
17	3.08	-4.40	-3.61	2.19
21	5.25	-4.70	-5.96	-3.17
25	3.37	7.11	7.28	9.86
29	11.06	10.07	-1.39	-10.09
33	-12.31	-8.52	-11.09	-10.03
37	-6.04	-5.51	-9.30	-10.33

```
41      -3.02        4.96       14.75        24.86
45      25.65       28.88      -14.15        -6.10
49      -9.92      -10.79       -6.77        -6.32
53      -4.83        -.08        1.87        -4.89
57      -7.04       -5.67      -10.64       -11.92
61     -10.12       -5.23       -4.09        -1.04
65       4.86       17.19       12.77        16.32
```

Intermediate values:

SUMS C2 = 17396.7, C*D = 8753.0, C*P = 10475.1
     D2 = 6126.9, D*P = 5499.2

Alpha1 = .151, Alpha2 = .898, Omega1 = .47, Omega2 = 2.52

Coefficient matrix:

```
34.206         .790        -.461         .426      124.829
  .790       33.794        -.454         .479       -7.036
 -.461        -.454       33.742        -.781      -33.500
  .426         .479        -.781       34.258       28.219
```

Solution vector:

```
3.635      -.317       -.930       .762
```

Trigonometric regression results using detrended data

```
 1       4.29         .84         .69       -1.20
 5      -3.72       -2.38       -4.57       -2.18
 9      -1.06        -.43        3.23        2.12
13       4.21        3.45        1.39        1.82
17      -2.17       -1.93       -3.36       -4.39
21      -1.80       -2.78         .67        1.72
25       2.23        4.76        2.31        3.24
29       1.05       -1.35       -1.09       -4.45
33      -2.90       -3.25       -2.67         .70
37        .12        3.49        3.41        2.97
41       3.91         .27         .59       -2.04
45      -3.65       -2.57       -4.56       -1.50
49      -1.06         .38        3.44        2.19
53       4.54        2.79        1.38        1.20
57      -2.70       -1.91       -3.91       -3.91
61      -1.68       -2.44        1.44        1.69
65       2.86        4.60        2.08        3.19
```

Final results with trend added back

```
 1      33.44       30.26       30.38       28.75
 5      26.50       28.12       26.19       28.85
 9      30.23       31.13       35.05       34.21
13      36.58       36.08       34.29       34.98
```

```
17      31.26      31.77      30.61      29.85
21      32.71      31.99      35.70      37.02
25      37.80      40.60      38.42      39.61
29      37.69      35.56      36.08      32.99
33      34.81      34.72      35.57      39.21
37      38.89      42.54      42.72      42.54
41      43.76      40.38      40.97      38.60
45      37.26      38.61      36.89      40.22
49      40.93      42.63      45.96      44.98
53      47.59      46.11      44.96      45.06
57      41.42      42.47      40.75      41.01
61      43.50      43.01      47.16      47.68
65      49.12      51.13      48.87      50.25
```

Bibliography

Aan, Peter W. "How RSI Behaves." *Futures* (January 1985).

Abraham, Bovas, and Johannes Ledolter. *Statistical Methods for Forecasting.* John Wiley & Sons, New York, 1983.

Acampora, Ralph, and Rosemarie Pavlick. "A Dow Theory Update." Originally published in the *MTA Journal* (January 1978), reprinted in the *MTA Journal* (Fall–Winter 2001).

Achelis, Steven B. *Technical Analysis from A to Z.* Probus, Chicago, 1995.

Ainsworth, Ralph M. *Profitable Grain Trading.* Traders Press, Greensville, SC, 1980.

Alexander, Carol. *Market Models, A Guide to Financial Data Analysis.* John Wiley & Sons, New York, 2001.

_____. *Market Risk Analysis, Volume 1, Quantitative Methods in Finance.* John Wiley & Sons, Hoboken, NJ, 2008 (2010 edition).

_____. *Market Risk Analysis, Volume 2, Practical Financial Econometrics.* John Wiley & Sons, Hoboken, NJ, 2008.

Andersen, Jon D. "Analyze Net Positions." *Technical Analysis of Stocks & Commodities* (July 2002).

Andrews, W. S. *Magic Squares and Cubes.* Dover, NY, 1960.

Angas, L. L. B. *Investment for Appreciation.* Somerset, NY, 1936.

Angell, George. "Thinking Contrarily." *Commodities Magazine* (November 1976). An interview with R. Earl Hadady.

_____. *Winning in the Commodities Markets.* Doubleday, New York, 1979.

Anon, Louis J. "Catch Short-Term Profits with ARIMA." *Commodities Magazine* (December 1981).

Appel, Gerald. *Winning Market Systems: 83 Ways to Beat the Market.* Signalert, Great Neck, NY, 1974.

_____. *Technical Analysis: Power Tool for Active Traders*, FT Prentice Hall, Upper Saddle River, NJ, 2005.

Appel, Gerald, and Martin E. Zweig. *New Directions in Technical Analysis.* Signalert, Great Neck, NY, 1976.

Arms, Richard W. Jr. "Equivolume—A New Method of Charting." *Commodities Magazine* (April 1973).

Arnold, Curtis M. *The Personal Computer Can Make You Rich in Stocks and Commodities.* Weiss Research, West Palm Beach, FL, 1980.

_____. "Tracking 'Big Money' May Tip Off Trend Changes." *Futures* (February 1985).

_____. "Your Computer Can Take You Beyond Charting." *Futures* (May 1984).

Aronson, David R. "An Introduction to Computer-Based Pattern Recognition." *Financial and Investment Software Review* (January/February 1984).

_____. "Artificial Intelligence/Pattern Recognition Applied to Forecasting Financial Market Trends." *Market Technicians Association Journal* (May 1985).

Arrington, George R. "Building a Variable-Length Moving Average." *Technical Analysis of Stocks & Commodities* (June 1991).

Aspray, Thomas. "Fine-Tuning the Demand Index." *Technical Analysis of Stocks & Commodities* (June 1986).

_____. "Demand Oscillator Momentum." *Technical Analysis of Stocks & Commodities* (September 1989).

Azoff, E. Michael. *Neural Network Time Series Forecasting of Financial Markets*. John Wiley & Sons, New York, 1994.

Barnett, Eugene H. *Programming Time—Shared Computers in BASIC*. Wiley-Interscience, New York, 1972.

Becker, John D. "Value of Oscillators in Determining Price Action." *Futures* (May 1994).

Bernstein, Jacob. "Cyclic and Seasonal Price Tendencies in Meat and Livestock Markets." In Todd Lofton, ed. *Trading Tactics: A Livestock Futures Anthology*. Chicago Mercantile Exchange, 1986.

_____. *The Handbook of Commodity Cycles*. John Wiley & Sons, New York, 1982.

_____. *Seasonal Concepts in Futures Trading*. John Wiley & Sons, New York, 1986.

_____. *Seasonality: Systems, Strategies, and Signals*. John Wiley & Sons, New York, 1998.

Bernstein, Peter L. *The Portable MBA in Investment*. John Wiley & Sons, New York, 1995.

Beyer, William H., ed. *Standard Mathematical Tables*, 24th ed. CRC Press, Cleveland, OH, 1976.

Bjorgen, Eric, and Steve Leuthold. "Corporate Insiders' Big Block Trades." *Journal of Technical Analysis* (Market Technicians Association, Winter–Spring 2002).

Blau, William. *Momentum, Direction, and Divergence*. John Wiley & Sons, New York, 1995.

Bollinger, John. *Bollinger on Bollinger Bands*. McGraw-Hill, New York, 2002.

_____. "John Bollinger of Bollinger Bands Fame." *Technical Analysis of Stocks & Commodities* (May 2002).

Bollinger, John A. Bollinger Capital Management, Inc. P.O. Box 3358, Manhattan Beach, CA 90266. www.bollingerbands.com.

Bolton, A. Hamilton. *The Elliott Wave Principle: A Critical Appraisal*. Bolton, Tremblay & Co., Montreal, 1960.

Bookstaber, Richard. *The Complete Investment Book*. Scott, Foresman, Glenview, IL, 1985.

Box, G. E. P., and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*, 2nd ed., Holden-Day, San Francisco, 1976.

Briese, Steve. "Commitment of Traders as a Sentiment Index." *Technical Analysis of Stocks & Commodities* (May 1990).

_____. "Tracking the Big Foot." *Futures* (March 1994).

_____. *The Commitment of Traders Bible*. John Wiley & Sons, Hoboken, NJ, 2008.

Brower, William. *Inside Edge*. Inside Edge Systems, 24 Old Stagecoach Road, Redding, CT 06896, March/April 1996.

Brown, Stephen J., and Mark P. Kritzman. *Quantitative Methods for Financial Analysis*, 2nd ed. Dow Jones-Irwin, Homewood, IL, 1990.

Brown, William G. S. "Logarithmic Point & Figure Charting." *Technical Analysis of Stocks & Commodities* (July 1995).

Bulkowski, Thomas, *The Encyclopedia of Chart Patterns*, John Wiley & Sons, New York, 2000.

_____. "What You Don't Know About Candlesticks." *Technical Analysis of Stocks & Commodities* (March 2011).

Burke, Gibbons. "Event-Based Analysis." *Futures* (April 1995).

Busch, Andrew. *World Event Trading.* John Wiley & Sons, Hoboken, NJ, 2007.

Carr, Michael J. "Get out the Vote and into Stocks." *Futures* (February 1996).

Cassidy, David. *Trading on Volume.* McGraw-Hill, New York, 2001.

Chande, Tushar S., and Stanley Kroll. *The New Technical Trader.* John Wiley & Sons, New York, 1994.

Chatfield, C. *The Analysis of a Time Series: Theory and Practice.* Chapman and Hall, London, 1975.

Chicago Board of Trade. *CBOT Market Profile.* 1984.

Church, A. H. *On the Relation of Phyllotaxis to Mechanical Laws.* Williams and Newgate, London, 1904.

Cleeton, Claud. *The Art of Independent Investing.* Prentice-Hall, Englewood Cliffs, NJ, 1976.

Cohen, A. W. *How to Use the Three-Point Reversal Method of Point and Figure Stock Market Trading.* Chartcraft, Larchmont, NY, 1972.

Colby, Robert W. *The Encyclopedia of Technical Market Indicators.* McGraw-Hill, New York, 2003.

Colby, Robert W., and Thomas A. Meyers. *The Encyclopedia of Technical Market Indicators.* Dow Jones-Irwin, Homewood, IL, 1988.

Commodity Trading Manual. Chicago Board of Trade.

Commodity Traders Club. *Comparative Performances.* Messena, NY, 1969. (A reprint).

Commodity Yearbook 1975. Commodity Research Bureau, New York, 1975.

Commodity Yearbook 1987. Commodity Research Bureau, New York, 1987.

Commodity Yearbook 1992. Commodity Research Bureau, New York, 1992.

Commodity Yearbook 1994. Knight-Ridder, New York, 1994.

CompuTrac Manual. Telerate, 1990.

Conners, Laurence A., and Blake E. Hayward. *Investment Secrets of a Hedge Fund Manager.* Probus, Chicago, 1995.

Conners, Laurence A., and Linda Bradford Raschke. *Street Smarts: High Probability Short-Term Trading Strategies.* M. Gordon Publishing Group, Malibu, CA, 1995.

Conners, Larry. "Fade the Breadth." *Futures* (January 2005)

_____. "Timing Your S&P Trades with the VIX." *Futures* (June 2002).

_____. *Short-Term Strategies That Work.* Conners, Jersey City, NJ, 2008.

ContiCommodity. *Seasonality in Agricultural Futures Markets.* ContiCommodity Services, Chicago, 1983.

Control Data Corporation. *Control Data 6000 Series Computer Systems Statistical Subroutines Reference Manual.* St. Paul, 1966.

Cootner, Paul, ed. *The Random Character of Stock Market Prices*, MIT Press, Cambridge, MA, 1964.

_____. "Speculation and Hedging." *Food Research Institute Studies*, *Vol. VII: 1967 Supplement.* Stanford University Press, Stanford, CA, 1967.

Covel, Michael W. *The Complete Turtle Trader.* Collins, New York, 2007.

Crabel, Toby. *Day Trading with Short-Term Price Patterns and Opening Range Breakout.* Traders Press, Greenville, SC 29603.

_____. "Opening Range Breakout, Part 4." *Technical Analysis of Stocks & Commodities* (February 1989).

Crim, Elias. "Are You Watching the 'Right' Signals?" *Futures* (June 1985).

CSI Technical Journal. Commodity Systems Inc., Boca Raton, FL (May 1995).

Cycles. Foundation for the Study of Cycles, Pittsburgh, January 1976.

Damodaran, Aswath. *Investment Valuation.* John Wiley & Sons, New York, 1996.

Davis, Jon. *Weather Sensitivity in the Markets.* Smith Barney, October 1994.

Davis, Robert Earl. *Profit and Profitability.* R. E. Davis, West LaFayette, IN, 1969.

Davis, R. E., and C. C. Thiel, Jr. *A Computer Analysis of the Moving Average Applied to Commodity Futures Trading.* Ouiatenon Management Co., West Lafayette, IN, 1970. (A research report.)

DeMark, Thomas R. "Retracing Your Steps." *Futures* (November 1995).

_____. *The New Science of Technical Analysis.* John Wiley & Sons, New York, 1994.

De Villiers, Victor. *The Point and Figure Method of Anticipating Stock Price Movements.* Trader Press, NY, 1966.

Dewey, Edward R., and Og Mandino. *Cycles.* Hawthorn Books, New York, 1971.

Donchian, Richard D. "Donchian's 5- and 20-Day Moving Averages." *Commodities Magazine* (December 1974).

Dorsey, Thomas. *Point & Figure Charting*, 2nd ed. John Wiley & Sons, New York, 2001.

Downie, N. M., and R. W. Heath. *Basic Statistical Method*, 3rd ed. Harper & Row, New York, 1970.

Duearte, Antonio Marcos, Jr., and Maria Luiza de Andrade Maia, "Optimal Portfolios with Derivatives," *Derivatives Quarterly* (Winter 1997).

Dunn & Hargitt Financial Services, Inc., West Lafayette, IN.

Dunn, Dennis. *Consistent Profits in June Live Beef Cattle.* Dunn & Hargitt, West Lafayette, IN, 1972.

Dunnigan, William. *One Way Formula.* Dunnigan, Palo Alto, CA, 1955.

_____. *Select Studies in Speculation.* Dunnigan, San Francisco 1954. (Includes "Gain in Grains" and "The Thrust Method in Stocks.")

_____. *117 Barometers for Forecasting Stock Price.* Dunnigan, San Francisco, 1954.

Earp, Richard B. "Correlating Taylor and Polous." *Commodities Magazine* (September 1973).

Edwards, Robert D., and John Magee. *Technical Analysis of Stock Trends.* John Magee, Springfield, MA, 1948.

Edwards, Robert D., John Magee, and W. H. Charles Bassetti. *Technical Analysis of Stock Trends*, 8th ed. Original edition published by John Magee.

Ehlers, John F. *Cybernetic Analysis for Stocks & Commodities.* John Wiley & Sons, Hoboken, NJ, 2004.

_____. "How to Use Maximum Entropy." *Technical Analysis of Stocks & Commodities* (November 1987).

_____. *MESA and Trading Market Cycles.* John Wiley & Sons, New York, 1992.

_____. *MESA and Trading Market Cycles*, 2nd ed. John Wiley & Sons, New York, 2001.

_____. "Moving Averages, Part 1." *Technical Analysis of Stocks & Commodities* (1988).

_____. "Moving Averages, Part 2." *Technical Analysis of Stocks & Commodities* (1988).

_____. "Optimizing RSI with Cycles." *Technical Analysis of Stocks & Commodities* (February 1986).

_____. "Relative Vigor Index." *Technical Analysis of Stocks & Commodities* (January 2002).

_____. *Rocket Science for Traders.* John Wiley & Sons, New York, 2001.

_____. "Trading Channels." *Technical Analysis of Stocks & Commodities* (April 1986).

Elder, Alexander. *Trading for a Living.* John Wiley & Sons, New York, 1993.

Elliott, R. N. *Nature's Law, The Secret of the Universe.* Elliott, New York, 1946.

_____. *The Wave Principle*. Elliot, New York, 1938.

Emmett, Tucker J. "Fibonacci Cycles." *Technical Analysis of Stocks & Commodities* (May 1983 and March/April 1984).

Eng, William F. "A Mechanical Trading System." *Technical Analysis of Stocks & Commodities* (July 1986).

_____. *The Technical Analysis of Stocks, Options & Futures*. Probus, Chicago, 1988.

Ericsson, Christina. *Forecasting Success*. Kaufman, Westport, CT, 1987.

Etzkorn, Mark. "All in a Day's Work." *Futures* (January 1995).

Evans, Eric. "Why You Can't Rely on 'Key Reversal Days.'" *Futures* (March 1985).

Faith, Curtis. *The Original Turtle Trading Rules* (2003), originalturtles.org.

Fischer, Robert. *The Golden Section Compass Seminar*. Fibonacci Trading, P.O. Box HM 1653, Hamilton 5, Bermuda.

Fisher, Andrew. "Teaching the Old Vix New Tricks." *Futures* (August 2001).

Fisher, Mark B. *The Logical Trader*. John Wiley & Sons, Hoboken, NJ, 2002.

Fishman, Mark B., Dean S. Barr, and Walter J. Loick. "Artificial Intelligence and Market Analysis." *Technical Analysis of Stocks & Commodities* (Bonus Issue 1993).

Floss, Carl William. *Market Rhythm*. Investors Publishing Co., New York, 1955.

Fosback, Norman G. *Stock Market Logic*. Dearborn Financial, Chicago, 1998.

_____. *Stock Market Logic*. The Institute for Economic Research, Ft. Lauderdale, FL, 1985.

Frederick, Randy. "Trading with Fair Value," *Futures* (April 2006)

Freeburg, Nelson F. *Formula Research*. Formula Research, Inc., 4745 Poplar Ave., Suite 307, Memphis, TN 38117-4408, Vol. VI, nos. 10–11, December 2000 and August 2001.

Fries, Christian P. "Elastic Moving Averages." *Technical Analysis of Stocks & Commodities* (June 2001).

Frost, A. J., and Robert R. Prechter, Jr. *Elliott Wave Principle*. New Classics Library, Chappaqua, NY, 1978.

Fuller, Wayne A. *Introduction to Statistical Time Series*. John Wiley & Sons, New York, 1976.

Fults, John Lee. *Magic Squares*. Open Court, La Salle, IL, 1974.

Futures Truth Magazine. Futures Truth, Hendersonville, NC. www.futurestruth.com.

Gann, William D. *The Basis of My Forecasting Method for Grain*, Lambert-Gann, Pomeroy, WA, 1976.

_____. *Forecasting Grains by Time Cycles*. Lambert-Gann, Pomeroy, WA, 1976.

_____. *Forecasting Rules for Cotton*. Lambert-Gann, Pomeroy, WA, 1976.

_____. *Forecasting Rules for Grain-Geometric Angles*. Lambert-Gann, Pomeroy, WA, 1976.

_____. *How to Make Profits in Commodities*. Lambert-Gann, Pomeroy, WA, 1976.

_____. *Master Calculator for Weekly Time Periods to Determine the Trend of Stocks and Commodities*. Lambert-Gann, Pomeroy, WA, 1976.

_____. *Master Charts*. Lambert-Gann, Pomeroy, WA, 1976.

_____. *Mechanical Method and Trend Indicator for Trading in Wheat, Corn, Rye, or Oats*. Lambert-Gann, Pomeroy, WA, 1976.

_____. *Rules for Trading in Soybeans, Corn, Wheat, Oats and Rye*. Lambert-Gann, Pomeroy, WA, 1976.

_____. *Speculation: A Profitable Profession (A Course of Instruction in Grains)*. Lambert-Gann, Pomeroy, WA, 1976.

_____. *45 Years in Wall Street*. Lambert-Gann, Pomeroy, WA, 1949.

_____. *Truth of the Stock Tape*. Available with *Wall Street Stock Selector*. Lambert-Gann, Pomeroy, WA, originally published in 1930.

_____. *Wall Street Stock Selector*. Available with *Truth of the Stock Tape*. Lambert-Gann, Pomeroy, WA, originally published in 1930.

Gannsoft Publishing. "Ganntrader I." *Technical Analysis of Stocks & Commodities* (January/February 1984).

Gately, Edward. *Neural Networks for Financial Forecasting*. John Wiley & Sons, New York, 1996.

Gehm, Fred. "Does Pyramiding Make Sense?" *Technical Analysis of Stocks & Commodities* (February 1986).

_____. *Commodity Market Money Management*. John Wiley & Sons, New York, 1983.

_____. *Quantitative Trading & Money Management*, rev. ed. Irwin, Homewood, IL, 1995.

Geman, Helyette. *Commodities and Commodity Derivatives*. John Wiley & Sons, Hoboken, NJ, 2005.

Gies, Joseph, and Frances Gies. *Leonard of Pisa and the New Mathematics of the Middle Ages*. Thomas M. Crowell, New York, 1969.

Gilchrist, Warren. *Statistical Forecasting*. John Wiley & Sons, London, 1976.

Ginter G., and J. Richie. "Data Errors and Price Distortions." In Perry J. Kaufman, ed. *Technical Analysis in Commodities*. John Wiley & Sons, New York, 1980.

Gotthelf, Edward B. *The Commodex System*. Commodity Futures Forecast, New York, 1970.

Gotthelf, Philip, and Carl Gropper. "Systems Do Work . . . But You Need a Plan." *Commodities Magazine* (April 1977).

Gould, Bruce G. *Dow Jones-Irwin Guide to Commodities Trading*. Dow Jones-Irwin, Homewood, IL, 1973.

Greer, Thomas V., B. Wade Brorsen, and Shi-Miin Liu. "Slippage Costs of a Large Technical Trader." *Technical Analysis of Stocks & Commodities* (January 1992).

Grushcow, J., and C. Smith. *Profits Through Seasonal Trading*. John Wiley & Sons, New York, 1980.

Gustafson, Gordon. "Price Persistency." *Technical Analysis of Stocks & Commodities* (January 2002).

Hadady, R. Earl. *Contrary Opinion*. Hadady, Pasadena CA, 1983.

_____. "Contrary Opinion." *Technical Analysis of Stocks & Commodities* (August 1988).

Hallberg, M. C., and V. I. West. *Patterns of Seasonal Price Variations for Illinois Farm Products*. Circular 861, University of Illinois College of Agriculture, Urbana, 1967.

Hambridge, Jay. *Dynamic Symmetry: The Greek Vase*. Yale University Press, New Haven, 1931.

_____. *Practical Applications of Dynamic Symmetry*. Yale University Press, New Haven, 1938.

Handmaker, David. "Low-Frequency Filters for Seasonal Analysis." In Perry J. Kaufman. *Handbook of Futures Markets*. John Wiley & Sons, New York, 1984.

_____. "Picking Software to Trade Technically." *Commodities* (August 1982).

Hannula, Hans. "Trading Planetary Eclipses." *Technical Analysis of Stocks & Commodities* (April 1992).

Harahus, David. ". . . *on Market Speculation* . . ." Harahus Analysis, Ferndale, MI, 1977.

Harris, Sunny, *TradeStation Made Easy*. John Wiley & Sons, Hoboken, NJ, 2011.

Haynie, F.M. "Doc." "Stretching the 'Profile' to Cover 24-Hour Markets." *Futures* (February 1992).

Haze, Van Court, Jr. *Systems Analysis: A Diagnostic Approach*. Harcourt, Brace & World, New York, 1967.

Hewison, Adam. "The Will Rogers Theory of Point & Figure Trading." *Technical Analysis of Stocks & Commodities* (August 1991).

Hieronymus, Thomas A. *Economics of Futures Trading*. Commodity Research Bureau, New York, 1971.

_____. *When to Sell Corn-Soybeans-Oats-Wheat.* University of Illinois College of Agriculture, Urbana, 1967.

Higgens, James E., and Allan M. Loosigian. "Foreign Exchange Futures." In Perry J. Kaufman, ed., *Handbook of Futures Markets.* John Wiley & Sons, New York, 1984.

Hildebrand, F. B. *Introduction to Numerical Analysis*, McGraw-Hill, New York, 1956.

Hill, Holliston. "Using Congestion Area Analysis to Set Up for Big Moves." *Futures* (April 1985).

Hill, John. "Simple vs. complex." *Futures* (March 1996).

Hirsch, Jeffrey A. and J. Taylor Brown. *The Almanac Investor: Profit from Market History and Seasonal Trends.* Stock Trader's Almanac, 2005.

Hochheimer, Frank L., and Richard J. Vaughn. *Computerized Trading Techniques 1982*. Merrill Lynch Commodities, New York, 1982.

Holland, John H. *Adaptations in Natural Language and Artificial Systems.* University of Michigan Press, Ann Arbor, 1975.

Holter, James T. "Delivering the Day-Trading Experience." *Futures* (August 2001).

Hoyle, *The Game in Wall Street and How To Play It Successfully.* Hoyle, originally published 1898, reprinted 1997.

"Hull Moving Average, The" *The Technical Analyst* (July–September 2010).

Hull, John C. *Fundamentals of Futures and Options Markets*, 4th ed. Prentice Hall, Upper Saddle River, NJ, 2002.

Hurst, J. M. *The Profit Magic of Stock Transaction Timing.* Prentice-Hall, Englewood Cliffs, NJ, 1970.

Hutson, Jack K. "Filter Price Data: Moving Averages vs. Exponential Moving Averages." *Technical Analysis of Stocks & Commodities* (May-June 1984).

_____. "Elements of Charting." *Technical Analysis of Stocks & Commodities* (March 1986).

_____. "Elements of Charting." *Technical Analysis of Stocks & Commodities* (March 1986).

_____. "Filter Price Data: Moving Averages Versus Exponential Moving Averages." *Technical Analysis of Stocks & Commodities* (May-June 1984).

_____. "Good TRIX." *Technical Analysis of Stocks & Commodities* (July 1983).

_____. "Using Fourier." *Technical Analysis of Stocks & Commodities* (January 1983).

Hutson, Jack K., and Anthony Warren. "Forecasting with Maximum Entropy." *Technical Analysis of Stocks & Commodities* (December 1984).

Jackson, J. T. *Detecting High-Profit Day Trades in the Futures Markets.* Windsor Books, Brightwaters, NY, 1994.

Jiler, William L. *Forecasting Commodity Prices with Vertical Line Charts.* Commodity Research Bureau, New York, 1966.

_____. *How Charts Are Used in Commodity Price Forecasting.* Commodity Research Publications, New York, 1977.

_____. *Volume and Open Interest: A Key to Commodity Price Forecasting.* Commodity Research Bureau, New York, 1967.

Johnson, A. Bruce. "Finding Cycles in Time Series Data." *Technical Analysis of Stocks & Commodities* (August 1990).

Johnson, Charles F. "Stochastic Oscillator Program for the HP-41C(V)." *Technical Analysis of Stocks & Commodities* (September-October 1984).

Jones, Donald. *Applications of the Market Profile.* CISCO, Chicago.

_____. "Locating Value with Auction Market Data." *Technical Analysis of Stocks & Commodities* (July 1989).

Kaeppel, Jay. "The January Barometer: Myth and Reality." *Technical Analysis of Stocks & Commodities* (July 1990).

_____. "The Stock Market, the Calendar, and You," *Technical Analysis of Stocks & Commoditie*s (December 2002).

_____. *Seasonal Stock Market Trends: The Definitive Guide to Calendar-Based Stock Market Trading*, John Wiley & Sons, Hoboken, NJ, 2009.

Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems." *Journal of Basic Engineering* (March 1960).

Kapner, Kenneth, and Robert McDonough. "Doing Your Homework on Individual Equity Futures." *Futures* (March 2002).

Kase, Cynthia. "Redefining Volatility and Position Risk." *Technical Analysis of Stocks & Commodities* (October 1993).

Katsanos, Markos. *Intermarket Trading Strategies*. John Wiley & Sons, Hoboken, NJ, 2008.

Kaufman, Perry J. *Alpha Trading*, John Wiley & Sons, Hoboken, NJ, 2011.

_____. *Handbook of Futures Markets.* John Wiley & Sons, New York, 1984.

_____. "High-Tech Trading." *Futures and Options World* (London, September 1985).

_____. "Market Momentum Re-Examined." Unpublished Article, November 1975.

_____. "Moving Averages and Trends." In Todd Lofton, ed. *Trading Tactics: A Livestock Futures Anthology.* Chicago Mercantile Exchange, 1986.

_____. "Price Shocks: Reevaluating Risk/Return Expectations." *Futures Industry* (June/July 1995).

_____. "Safety-Adjusted Performance Evaluation." *Journal of Futures Markets*, 1 (1981).

_____. *A Short Course in Technical Trading.* John Wiley & Sons, Hoboken, NJ, 2003.

_____. *Smarter Trading.* McGraw-Hill, New York, 1996.

_____. "Technical Analysis." In Nancy H. Rothstein, ed. *The Handbook of Financial Futures.* McGraw-Hill, New York, 1984.

_____. *Technical Analysis in Commodities.* John Wiley & Sons, New York, 1980.

Kaufman, Perry J., Dennis C. Koutras, and Peter Psycharis. "Positioning the Greek Equities Market in International Portfolios." *Nonlinear Analysis, Theory, Methods & Applications* 30, no. 4. *Proceedings of the 2nd World Congress of Nonlinear Analysts*, Elsevier Science Ltd., Great Britain 1997.

Kaufman, Perry, and Alberto Vivanti. *Global Equity Investing.* McGraw-Hill, New York, 1997.

Kaufman, Perry J., and Kermit C. Zieg, Jr. "Measuring Market Movement." *Commodities Magazine* (May 1974).

Kelly, John L., Jr. "Kelly Betting System." *Bell System Technical Journal* (July 1956).

Keltner, Chester W. *How to Make Money in Commodities.* The Keltner Statistical Service, Kansas City, 1960.

Kemeny, John G., and J. Laurie Snell. *Finite Markov Chains.* Springer-Verlag, New York, 1976.

Kemeny, John G., and Thomas E. Kurtz. *Basic Programming*, 2nd ed. John Wiley & Sons, New York, 1971.

Kepka, John F. "Trading with ARIMA Forecasts." *Technical Analysis of Stocks & Commodities* (August 1985).

Kestner, Lars. *Quantitative Trading Strategies*. McGraw-Hill, New York, 2003.

Klein, Frederick C., and John A. Presbo. *News and the Market.* Henry Regnery Co., Chicago, 1974.

Knapp, Volker. "Parabolic Alternatives," *Active Trader Magazine* (September 2010).

Knuth, Donald E. *The Art of Computer Programming*, *Vol. 2: Seminumeric Algorithms.* Addison-Wesley, Reading, MA, 1971.

Koy, Kevin. *Markets 101.* MLS Publishing, 401 S. LaSalle St., Chicago, IL 60605.

Kraska, Don. "Letters to S&C." *Technical Analysis of Stocks & Commodities* (February 1996).

Krausz, Robert. Fibonacci Trader, *Tutorial & Educational Supplement.* Fibonacci Trader Corporation, 757 S.E. 17th Street, Suite 272, Ft. Lauderdale, FL 33316.

_____. "Intraday Strategies for Multiple Time Frames." *Futures* (January 2002).

_____. "A Strategy for Trading Multiple Time Frames." *Futures* (November 2001).

_____. *W.D. Gann Treasure Discovered.* Geometric Traders Institute, Fibonacci Trader Corp, 450-106 State Road 13 North, #206, Jacksonville, FL 32259-3863.

Kritzman, Mark P., and Stephen J. Brown. *Quantitative Methods for Financial Analysis*, 2nd ed. Dow Jones-Irwin, Homewood, IL, 1990.

Kritzman, Mark, and Don Rich. "The Mismeasurement of Risk." *Financial Analysts Journal* (May-June 2002).

Kroll, Stanley. *The Professional Commodity Trader.* Harper & Row, New York, 1974.

Kroll, Stanley, and Irwin Shishko. *The Commodity Futures Market Guide*, Harper & Row, New York, 1973.

Kudyba, Stephan. "Are neural networks a better forecaster?" *Futures* (October 1998).

Kunz, Kaiser S. *Numerical Analysis*, McGraw-Hill, New York, 1957.

Labys, Walter C. *Dynamic Commodity Models: Specification, Estimation, and Simulation.* Lexington Books, Lexington, MA, 1973.

Lafferty, Patrick E. "End-Point Moving Average." *Technical Analyisis of Stocks & Commodities* (October 1995).

Lambert, Donald R. "Commodity Channel Index: Tool for Trading Cyclic Trends." *Commodities* (1980). Reprinted in *Technical Analysis of Stocks & Commodities* (July 1983).

_____. "Exponentially Smoothed Moving Averages." *Technical Analysis of Stocks & Commodities* (September-October 1984).

_____. "The Market Directional Indicator." *Technical Analysis of Stocks & Commodities* (November-December 1983).

Lane, George C. "Lane's Stochastics." *Technical Analysis of Stocks & Commodities* (May/June 1984).

Larson, Arnold. "Measurement of a Random Process in Futures Pricing." *Food Research Institute Studies* 1, no. 3 (November 1960).

Lee, Robert T. H. *Power Tools for Traders.* Mega Capital Limited, Hong Kong, 1997.

Lefèvre, Edwin. *Reminiscences of a Stock Operator.* Books of Wall Street, Burlington, VT, 1980. Originally published in 1923 by George H. Doran Co., New York.

Lemay, Phillippe. http://tecfa.unige.ch.

Levy, David. *Computer Chess Compendium.* Reprint by Ishi Press, New York, 2009.

Levy, David and Monty Newborn. *How Computers Play Chess.* Computer Science Press, New York, 1991.

Lincoln, Thomas H. "Time Series Forecasting: ARMAX." *Technical Analysis of Stocks & Commodities* (September 1991).

Little, Jeffrey B., and Lucien Rhodes. *Understanding Wall Street*, 3rd ed. McGraw-Hill, New York, 1991.

Lo, Andrew W., Harry Mamaysky, and Jiang Wang. "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation." *Journal of Finance* (August 2000).

Lofton, Todd. "Chartists Corner." *Commodities Magazine* (December 1974). Two series of articles.

_____.*Getting Started in Futures*, 4th ed. John Wiley & Sons, New York, 2001.

_____. "Moonlight Sonata." *Commodities Magazine* (July 1974).

_____. *Trading Tactics: A Livestock Futures Anthology.* Chicago Mercantile Exchange, 1986.

Long, Jeanne. "Planetary Support and Resistance on the DJIA." *A Traders Astrological Almanac 1994.* Professional Astrology Service, 757 S.E. 17th St., Ft. Lauderdale, FL 33316.

Longerstaey, Jacques. *RiskMetrics' Technical Document*, 4th ed. Morgan Guaranty Trust Company, riskmetrics@jpmorgan.com.

Lorca [Susino], Francisco J. "Exploiting Stock Market Cycles," *Futures* (April 2009).

Loyo, Luis Ballesca. "Price Projections on Point and Figure Charts." *Technical Analysis of Stocks & Commodities* (July 1989).

Luce, R. Duncan, and Howard Raiffa. *Games and Decisions.* John Wiley & Sons, New York, 1957.

Lukac, Louis P., B. Wade Brorsen, and Scott H. Irwin. "How to Test Profitability of Technical Trading Systems." *Futures* (October 1987).

_____. *Similarity of Computer Guided Technical Trading Systems.* CSFM-124, Working Paper Series, Columbia Futures Center, Columbia University Business School, New York, March 1986.

Maccaro, James. "The Early Chartists: Schabacker, Edwards, Magee." *Technical Analysis of Stocks & Commodities* (November 2002).

MacKay, Charles. *Extraordinary Popular Delusions and the Madness of Crowds.* Noonday Press (Farrar, Straus, and Giroux), New York, 1932.

Macon, Nathaniel. *Numerical Analysis*, John Wiley & Sons, New York, 1963.

Madlem, Peter. *Power Investing with Sector Funds.* St. Lucie Press/American Management Association, Del Ray Beach, FL, 1999.

Mandelbrot, B. B. "A Multifractional Walk Down Wall Street." *Scientific American* (February 1999).

Markowitz, Harry M. "Portfolio Selection." *Journal of Finance* (1979).

_____. *Portfolio Selection.* Cowles Foundation, Yale University Press, New Haven, 1959.

Mart, Donald S. *The Master Trading Formula.* Winsor Books, Brightwaters, NY, 1981.

Martin, Francis F. *Computer Modeling and Simulation.* John Wiley & Sons, New York, 1968.

Maxwell, Joseph R. St. *Commodity Futures Trading with Moving Averages.* Speer, Santa Clara, CA, 1974.

McGinley, John. *Technical Trends.* P.O. Box 792, Wilton, CT 06897.

McKallip, Curtis, Jr. "Fundamentals Behind Technical Analysis." *Technical Analysis of Stocks & Commodities* (November 1989).

McKenna, Peter. *The Event-Trading Phenomena.* TradeWins, 2005.

McKinsey, J. C. C. *Introduction to the Theory of Games.* McGraw-Hill, New York, 1952.

McMillan, Lawrence G. "Put-Call Ratios." *Technical Analysis of Stocks & Commodities* (October 1995).

McNicholl, Dennis. "Better Bollinger Bands." *Futures* (October 1998).

_____. "Old Statistical Methods for New Tricks in Analysis." *Futures* (April 2002).

McNutt, Malcolm. "First-Hour Breakout System." *Technical Analysis of Stocks & Commodities* (July 1994).

Meeus, Jean. *Astronomical Algorithms.* Willman-Bell, Inc., P.O. Box 35025, Richmond, VA 23235.

_____. *Astronomical Formulae for Calculators*. Willman-Bell, Inc., P.O. Box 35025, Richmond, VA 23235.

Meissner, Gunter, and Sarp Cercioglu. "Trading Trends and Correlations." *Futures* (January 2003).

Mendenhall, William, and James E. Reinmuth. *Statistics for Management and Economics*, 2nd ed. Duxbury Press, North Scituate, MA, 1974.

Merrill, Arthur A. *Behavior of Prices on Wall Street*. The Analysis Press, Chappaqua, NY, 1966.

_____. "The Presidential Election Cycle." *Technical Analysis of Stocks & Commodities* (March 1992).

Meyers, Dennis. "Range Roving." *Active Trader* (March 2003).

Mills, Frederick Cecil. *Statistical Methods*. Henry Holt & Co., New York, 1924.

Mills, Mark. "Regularization" (in "Traders' Tips"). *Technical Analysis of Stocks & Commodities* (July 2003).

Montgomery, Douglas C., and Lynwood A. Johnson. *Forecasting and Time Series*. McGraw-Hill, New York, 1976.

Moore, Steve. "Playing the Seasonals." *Active Trader* 1, no. 8 (November 2000).

Morney, M. J. "On the Average and Scatter." In *The World of Mathematics*, Vol. 3 (James R. Newman, ed.). Simon & Schuster, New York, 1956.

Murphy, John. *Technical Analysis of the Futures Markets*, 2nd ed. New York Institute of Finance, 2001.

Murphy, John J. *Intermarket Technical Analysis*, John Wiley & Sons, New York, 1991.

Narang, Rishi K. *Inside the Black Box*, John Wiley & Sons, Hoboken, NJ, 2009.

Narcouzi, Christopher. "Winning with Failures." *Technical Analysis of Stocks & Commodities* (November 2001).

Neill, Humphrey. *The Art of Contrary Thinking*. The Caxton Printers, Caldwell, OH, 1960.

Noble, Grant. "The Best Trading Indicator—the Media." *Technical Analysis of Stocks & Commodities* (October 1989).

Nofri, Eugene. *The Congestion Phase System*. A republication is available on successincommodities .com (2010)

O'Higgins, Michael. *Beating the Dow*. HarperCollins, 1991.

Omega TradeStation. *Easy Language Manual*. Omega Research, Boca Raton, FL 1995.

Oster, Carol. "Support for Resistance: Technical Analysis and Intraday Exchange Rates." *FRBNY Economic Policy Review* (July 2000).

Oster, Merrill J. "How the Young Millionaires Trade Commodities." *Commodities Magazine* (March and April 1976).

Oster, Merrill, J., et al. *How to Multiply Your Money*. Investors Publications, Cedar Falls, IA, 1979.

Pardo, Robert. *Design, Testing and Optimization of Trading Systems*. John Wiley & Sons, New York, 1992.

Parker, Derek, and Julia Paricor. *The Compleat Astrologer*. McGraw-Hill, New York, 1971.

Payne, J. J. "A Better Way to Smooth Data." *Technical Analysis of Stocks & Commodities* (October 1989).

Perrine, Jack. "Taurus the Bullish." *Commodities Magazine* (September 1974).

Person, John L. "Pivot Points and Candles." *Futures* (February 2003).

Peters, Edgar. *Chaos and Order in the Capital Markets*. John Wiley & Sons, New York, 1991.

Peterson, Dennis D. "Positive Volume Index." *Technical Analysis of Stocks & Commodities* (April 2003).

Picerno, James. *Dynamic Asset Allocation*. Bloomberg, New York, 2010.

Pisani, Robert. "How Market Structure Helps You Analyze Price Behavior." *Futures* (October 1987).

Polous, E. Michael. "The Moving Average as a Trading Tool." *Commodities Magazine* (September 1973).

Powers, Mark J. *Getting Started in Commodity Futures Trading.* Investors Publications, Waterloo, Iowa, 1975.

Prathap, Johnan. "Three-Bar Inside Bar Pattern," *Technical Analysis of Stocks & Commodities* (March 2011).

Prechter, Robert R., Jr. "Computerizing Elliott." *Technical Analysis of Stocks & Commodities* (July 1983).

_____. "Major Sea Change II." *Futures* (March 1996).

_____. *The Major Works of R. N. Elliott*, New Classics Library, Chappaqua, NY, 1980.

Prechter, Robert R., Jr., David Weiss, and David Allman. "Forecasting with the Elliott Wave Principle." In Todd Lofton, ed. *Trading Tactics: A Livestock Futures Anthology.* Chicago Mercantile Exchange, 1986.

Pring, Martin J. *Martin Pring on Market Momentum.* International Institute for Economic Research, Gloucester, VA, 1993.

_____. *Technical Analysis Explained.* McGraw-Hill, New York,1991.

_____. *Trading Systems Explained.* Marketplace Books, Columbia, MD 2003.

_____. "Twice as Nice: The Two-Bar Reversal Pattern." *Active Trader* (March 2003).

Reiman, Ray. "Handicapping the Grains." *Commodities Magazine* (April 1975).

Reverre, Stephane. "Trading Against the Gap." *Technical Analysis of Stocks & Commodities* (Bonus Issue 2003).

_____. "Trading the Opening Gap." *Technical Analysis of Stocks & Commodities* (November 1999).

Rockwell, Charles S. "Normal Backwardation, Forecasting, and the Returns to Commodity Futures Traders." *Food Research Institute Studies*, *Vol. VII: 1967 Supplement.* Stanford University Press, Stanford, CA, 1967.

Ross, Joe. *Trading Spreads and Seasonals.* Ross Trading, 1995.

Rotella, Robert P. *The Elements of Successful Trading.* The New York Institute of Finance, 1992.

Rothstein, Nancy H. *The Handbook of Financial Futures.* McGraw-Hill, New York, 1984.

Ruckdeschel, F. R. *BASIC Scientific Subroutines*, Vol. 1. Byte/McGraw-Hill, Peterborough, NH, 1981.

_____. *BASIC Scientific Subroutines*, Vol. 2. Byte/McGraw-Hill, Peterborough, NH, 1983.

Rudd, Barry. *Stock Patterns for Day Trading and Swing Trading.* Traders Press, Greenville, SC, 1998.

Ruggiero, Murray A., Jr. "Artificial Trader Jumps Candlesticks." *Futures* (February 1995).

_____. "Build a Real Neural Net." *Futures* (June 1996).

_____. "Building the Wave." *Futures* (April 1996).

_____. *Cybernetic Trading Strategies.* John Wiley & Sons, New York, 1997.

_____. "Exploiting Report Day Tendencies in Treasuries." *Futures* (February 2001).

_____. "Lighting Candlesticks with Fuzzy Logic." *Futures* (March 1996).

_____. "The Science of Pattern-Based Trading." *Futures* (2004).

_____. "Seeking a Commitment." *Futures* (April 2002).

Saitta, Alex. "A Price and Volume-Based System." *Technical Analysis of Stocks & Commodities* (March 1996).

Salov, Vlarii. "Market Profile and the Distribution of Prices." *Futures* (June 2011).

Schabacker, R. W. *Stock Market Theory and Practice.* B. C. Forbes, New York, 1930.

Schirding, Harry. "Stochastic Oscillator." *Technical Analysis of Stocks & Commodities* (May-June 1984).

Schwager, Jack D. *A Complete Guide to the Futures Markets.* John Wiley & Sons, New York, 1984.

_____. *Risk and Money Management.* In Perry J. Kaufman, ed. *Handbook of Futures Markets.* John Wiley & Sons, New York, 1984.

Schwager, J. D. "A Trader's Guide to Analyzing the Potato Futures Market." *1981 Commodity Yearbook.* Commodity Research Bureau, New York, 1981.

Schwager, Jack D. *Schwager on Futures: Fundamental Analysis.* John Wiley & Sons, New York, 1995.

_____. *Schwager on Futures: Technical Analysis.* John Wiley & Sons, New York, 1996.

Scorpio, Massimiliano. "Targeting Your Pattern," *Technical Analysis of Stocks & Commodities* (July 2005).

Seidel, Andrew D., and Philip M. Ginsberg. *Commodities Trading.* Prentice-Hall, Englewood Cliffs, NJ, 1983.

Sipley, Richard. Market Indicators. Bloomberg, New York, 2009.

Shaleen, Kenneth H. *Volume and Open Interest.* Probus, Chicago, 1991.

Shaw, John E. B. *A Professional Guide to Commodity Speculation.* Parker Publ., West Nyack, NY, 1972.

Sheard, Robert. *The Unemotional Investor.* Fireside, New York, 1998.

Sklarew, Arthur. *Techniques of a Professional Commodity Chart Analyst.* Commodity Research Bureau, New York, 1980.

Smith, Adam. *The Money Game.* Random House, New York, 1967.

Springer, Clifford H., Robert E. Herlihy, and Robert I. Beggs. *Advanced Methods and Models.* Richard D. Irwin, Homewood, IL, 1965.

_____. *Probabilistic Models.* Richard D. Irwin, Homewood, IL, 1968.

Statistical Annual. Chicago Board of Trade, Chicago, 1969–1975 eds.

Steidlmayer, J. Peter. *Steidlmayer on Markets.* John Wiley & Sons, New York, 1989.

Steidlmayer, J. Peter, and Kevin Koy. *Markets & Market Logic.* The Porcupine Press, Chicago, 1986.

Steidlmayer, J. Peter, and Shera Buyer. *Taking the Data Forward.* Market Logic Inc., Chicago, 1986.

Steinberg, Jeanette Nofri. "Timing Market Entry and Exit." *Commodities Magazine* (September 1975).

Stigum, Marsha. *Money Market Calculations.* Dow Jones-Irwin, Homewood, IL, 1981.

Strahm, Norman D. "Preference Space Evaluation of Trading System Performance." In Perry J. Kaufman, ed. *Handbook of Futures Markets.* John Wiley & Sons, New York, 1984.

Stridsman, Thomas. "Revelation Trading," *Futures* (February 1998).

Sweeney, John. *Campaign Trading.* John Wiley & Sons, New York, 1997.

_____. "Maximum Adverse Excursions for Stops." *Technical Analysis of Stocks & Commodities* (April 1987).

Taylor, George Douglass. *The Taylor Trading Technique.* Lilly Pub. Co., Los Angeles, 1950. (This book has been reprinted by Traders Press, P.O. Box 6206, Greenville, SC 29606 in 1994.)

Taylor, Robert Joel. "The Major Price Trend Directional Indicator." *Commodities Magazine* (April 1972).

Taylor, William T. "Fourier Spectral Analysis." *Technical Analysis of Stocks & Commodities* (July-August 1984).

Teleb, Nassim Nicholas. *Fooled by Randomness.* Texere, New York, 2001.

Tesco, Rudy. "LRS + R-Squared = The 95% Solution." *Technical Analysis of Stocks & Commodities* (February 2003).

Teweles, Richard J., Charles V. Harlow, and Herbert L. Stone. *The Commodity Futures Game, Who Wins? Who Uses? Why?* McGraw-Hill, New York, 1974.

The Traders Note Book. Dunn & Hargitt Financial Services, W. Lafayette, IN, 1970.

Thiel, Charles, and R. E. Davis. *Point and Figure Commodity Trading: A Computer Evaluation.* Dunn & Hatgitt, West LaFayette, IN, 1970.

Thiriea, H., and S. Zionts, eds. *Multiple Criteria Decision Making.* Springer-Verlag, Berlin, 1976.

Thompson, Jesse H. "The Livermore System." *Technical Analysis of Stocks & Commodities* (May 1983).

_____. "What Textbooks Never Tell You." *Technical Analysis of Stocks & Commodities* (November-December 1983).

Thorp, Edward O. *Beat the Dealer.* Vintage, New York, 1966.

Tippett, L. C. "Sampling and Standard Error." *The World of Mathematics*, Vol. 3 (James R. Newman, ed.). Simon & Schuster, New York, 1956.

Townsend, Robert. *Up the Organization.* Knopf, New York, 1970.

Trading Strategies. Futures Symposium International, Tucson, AZ, 1984.

Tubbs, Frank. "Tubbs' Stock Market Correspondence Lessons." (Chap. 13, Tape Reading.)

Turner, Dennis, and Stephen H. Blinn. *Trading Silver—Profitably.* Arlington House, New Rochelle, NY, 1975.

Vince, Ralph. *Portfolio Management Formulas.* John Wiley & Sons, New York, 1990.

_____. *The Handbook of Portfolio Mathematics*, John Wiley & Sons, Hoboken, NJ, 2007.

Visual Numerics. *IMSL Math Library.* 1997.

Van Horne, James C. *Financial Market Rates and Flows*, 6th ed. Prentice Hall, Upper Saddle River, NJ, 2001.

von Neumann, John, and Oskar Morgenstern. *Theory of Games and Economic Behavior.* 1943; Princeton University Press, Princeton, 1953.

von Ronik, Wolf. "System Testing via Monte Carlo." *Futures* (April 2001).

Walker, Jeff. "What K-wave?" *Technical Analysis of Stocks & Commodities* (July 1990).

Wall, E. Marshall. "Rolling with the Punches." *Futures* (July 1996).

Warwick, Ben. *Event Trading.* Irwin, Homewood, IL, 1996.

Warren, Anthony. "A Mini Guide to Fourier Spectrum Analysis." *Technical Analysis of Stocks & Commodities* (January 1983).

_____. "Fourier Analysis! In a Nutshell; Faster and Better." *Technical Analysis of Stocks & Commodities* (December 1983).

_____. "An Introduction to Maximum Entropy Method (MEM) Technical Analysis." *Technical Analysis of Stocks & Commodities* (February 1984).

_____. "Optimizing the Maximum Entropy Method." *Technical Analysis of Stocks & Commodities* (March-April 1984).

Warren, Anthony, and Jack K. Hutson. "Finite Impulse Response Filter." *Technical Analysis of Stocks & Commodities* (May 1983).

Waters, James, J., and Larry Williams. "Measuring Market Momentum." *Commodities Magazine* (October 1972).

Watling, T. F., and J. Morley. *Successful Commodity Futures Trading.* Business Books Ltd., London, 1974.

Watson, Donald S., and Mary A. Holman. *Price Theory and Its Uses*, 4th edition. Houghton Mifflin, Boston, 1977.

Waxenberg, Howard K. "Technical Analysis of Volume." *Technical Analysis of Stocks & Commodities* (March 1986).

Weiss, Eric. "Applying ARIMA Forecasts." *Technical Analysis of Stocks & Commodities* (May 1983).

_____. "ARIMA Forecasting." *Technical Analysis of Stocks & Commodities* (January 1983).

Weymar, F. H. *The Dynamics of the World Cocoa Market.* MIT Press, Cambridge, MA, 1968.

White, Adam. "The Eight-Year Presidential Election Pattern." *Technical Analysis of Stocks & Commodities* (November 1994).

Wilder, J. Welles, Jr. *Chart Trading Workshop 1980.* Trend Research, Greensboro, NC, 1980.

_____. *New Concepts in Technical Trading*, Trend Research, Greensboro, NC, 1978.

Williams, Bill. *Trading Chaos.* John Wiley & Sons, New York, 1995.

Williams, Billy. "Biting off profits with the Rattlesnake Breakout Method." *Futures* (October 2010).

Williams, Edward E., and M. Chapman Findlay III. *Investment Analysis.* Prentice-Hall, Englewood Cliffs, NJ, 1974.

Williams, J. D. *The Complete Strategyst.* McGraw-Hill, New York, 1966.

Williams, Larry. "The Ultimate Oscillator." *Technical Analysis of Stocks & Commodities* (August 1985).

Williams, Larry R. *How I Made One Million Dollars … Last Year … Trading Commodities.* Conceptual Management, Carmel Valley, CA, 1973.

Williams, Larry, and Charles Lindsey. *The Trident System.* Lindsey, O'Brien & Co., Thousand Oaks, CA, 1975. (A report.)

Williams, Larry, and Michelle Noseworthy. "How Seasonal Influences Can Help You Trade Commodities." *Commodities Magazine* (October 1976).

Winski, Joseph N. "A Sure Thing?" *The Dow Jones Commodities Handbook 1977.* Dow-Jones Books, Princeton, NJ, 1977.

Wolberg, John R. *Expert Trading Systems.* John Wiley & Sons, New York, 2000.

Working, Holbrook. "Test of a Theory Concerning Floor Trading on Commodity Exchanges." *Food Research Institute Studies, Vol. VII: 1967 Supplement.* Stanford University Press, Stanford, CA, 1967.

Wyckoff, Richard D. *The Richard D. Wyckoff Method of Trading and Investing in Stocks.* Wyckoff Associates, Inc., Park Ridge, IL, 1936.

_____. *Stock Market Technique, Number One.* Wyckoff, NY, 1933.

_____. *Wall Street Ventures and Adventures Through Forty Years.* Harper & Brothers, New York, 1930.

Xin Xie. Full View Integrated Technical Analysis. John Wiley & Sons (Asia), 2011.

Zieg, Kermit C., Jr., and Perry J. Kaufman. *Point and Figure Commodity Trading Techniques.* Investors Intelligence, Larchmont, NY, 1975.

About the Companion Website

This book is accompanied by a website, which you can find at www.wiley.com/go/tradingsystems5e (password: pjk123). The site includes TradeStation (version 9) code broken up by category (indicators, strategies, and functions) and Excel 2010 spreadsheets. The Companion Website provides a list of the contents in two ways: alphabetically and by chapter. For readers not familiar with TradeStation and EasyLanguage, I recommend Sunny Harris' *TradeStation Made Easy*. It is expected that MetaStock code will be added to the website in the near future.

 Note: The items with an asterisk (*) are not referenced directly the book, but are part of the Companion Website.

Index