System Design and Testing

From Charles D. Kirkpatrick II and Julie R. Dahlquist, *Technical Analysis: The Complete Resource for Financial Market Technicians*, 3rd Edition (Old Tappan, New Jersey: Pearson Education, Inc., 2016), Chapter 22.

Learning Objective Statements

- Assess the value and challenges of using a system for trading or investing
- Compare and analyze differences between discretionary and nondiscretionary systems
- Evaluate the mind-set and discipline required to develop and trade with a system
- Organize the basic procedures for designing a system
- Inventory types of technical trading systems
- Defend the necessity of risk management protocols in a trading system
- Examine critical aspects of performing system tests
- Compare and evaluate standard measures of system profitability and risk
- Differentiate between various methods of optimization

We have covered most of the methods used by technical analysts to analyze the trading markets. We now enter into the field of how to utilize this knowledge to produce profits and reduce risk. Any constant and consistent management of assets requires money management and some type of system. Haphazardly investing or trading on intuition, rumor, or untested theories is a road to disaster. It is why most amateur traders and investors lose money.

A fundamental investor may use price-to-earnings, debt ratios, and so forth, whereas a technical investor will likely use relative strength, price trend, or volatility, and both investors will believe they are doing the correct analysis. They are not. In both cases, the methods may be correct, but making money requires a tested system. There are many myths in investing, and most investors succumb to them without further analysis. To trade or invest successfully, we need to know not only how profitable a method has been, but also what the risks of capital loss were. Not having an understanding of tested methods is flying blind in the financial markets. How do we test these methods? We create a system. The system must include not only the method for profit but also the means of controlling risk of loss. Both aspects of investing are extremely important. Some would argue that controlling loss is even more important than the profit method, that by buying and selling at the flip of a coin, one could make a decent return just by controlling risk of loss.

Let us begin this chapter by looking a little more closely at what a system is. Then we focus on what risk is and how to control losses. Once we have these foundations, we can focus on the mechanics of developing a system and testing investment strategies.

Why Are Systems Necessary?

No stock market goes up forever. Indeed, most world stock markets have declined to zero at one time or another. The buy-and-hold strategy so popular in the United States today is based on a statistical anomaly. It is a strategy based on a survival bias in the U.S. and the U.K. markets, the only countries in history, so far, whose markets have not completely disappeared at some time (Burnham, 2005). This has caused a misleading assumption that U.S. stocks and stocks in general will necessarily continue to rise. "It would be naïve to expect the future of U.S. stocks to be as bright as the past" (Burnham, 2005, p. 175).

We certainly know that individual stocks can go to zero. How about buggy whips in 1910, or canals in 1830, or bowling in 1950, or junk bonds and REITs in 1980, or more recently the autos and the banks? Thus, a long-term plan that excludes a means of controlling risk is eventually doomed.

On the other hand, most technical and fundamental methods, by themselves, are not profitable over time either. Some of the exceptions have been covered earlier in this book, but these methods primarily depend upon the market circumstances at the time, on the method used, and on controlling risk. Traders' and investors' greatest misconception is that the market has order and that by finding and acting on that order, profits will be consistent and large. It presumes that a magic formula exists somewhere that can predict markets. This belief is not true. In looking at the previous studies in this book, there is no magic order to the markets beyond the fact that they sometimes trend and, more often, remain in trading ranges. The money made is based on the use of well-controlled entries and exits, especially those that limit the amount of loss that can occur and that will react to changing conditions in the market. A system will aid the investor or trader in timing these market entries and exits.

Discretionary Versus Nondiscretionary Systems

Systems are the next step in the development of an investment plan after understanding the methods of either technical or fundamental investing. Systems can be discretionary, nondiscretionary, or a combination of both. In discretionary systems, entries and exits are determined by intuition; in other words, the trader or investor exercises some discretion in making trades. Nondiscretionary systems are those in which entries and exits are determined mechanically by a computer.

Think for a minute of the stereotypical discretionary trader. Imagine the ultimate discretionary trader behaving like the man in the antacid advertisement with two or three phones yelling, "Buy" in one, "Sell" in another, with computer screens showing prices and charts of securities all over the world, with ringing phones, with news broadcasts from financial TV stations, and with a large contact list of people in different specialties. This type of trader is generally looking for the home run. It is a great image, one that has in it a bit of the swashbuckler, the gunslinger, and so on. In fact, many truly exceptional traders are like this. They have the gifted intuition to be able to do this consistently and profitably.

Most people, however, do not have the time, the knowledge, the contacts, the equipment, the quickness of thought, or the stomach to do this. In fact, most people who attempt to trade like this either burn out or go broke. They have no way of evaluating what they are doing except from the equity in their account at the end of the day. It is as if the excitement is more important than making profits.

The nondiscretionary trader, on the other hand, is usually calm, calculating, and likely bored. The majority of successful traders and investors use nondiscretionary systems (Etzkorn interview of Babcock, 1996). Some have been engineers; others have that type of mind, familiar with statistics and systems. They have studied the markets, the methods of profit-making—both fundamental and technical—and have tested the techniques using modern statistical methods. They understand that nothing is perfect and that markets change character over time. However, by testing their methods and strategies, they have derived a mechanical system that minimizes risk of loss and maximizes return.

Rules are the structure of a system. An example of a rule would be "buy when one moving average A crosses above another moving average B." Variables are the numerical inputs required in the rules (length of moving average A and length of moving average B). Parameters are the actual values used in the variables (two days and seven days). A system will include all these factors; their usefulness is determined by testing different rules, variables, and parameters over varied markets and market conditions.

A purely nondiscretionary system is one that runs by itself on market data that is continually fed into it. If our rule is to buy when the two-day moving average crosses above the seven-day moving average, for example, a buy order automatically is placed when this occurs. Once the trader has determined the rule to follow, the system is on autopilot and the trader does not make decisions.

A trader or investor can also choose to use a partial discretionary system. The partial discretionary system is one that generates signals that then are acted upon by the investor based on personal confidence in them and experience with them. By having some discretion, however, the system cannot be tested accurately because emotion can enter into the trading decisions and cause unquantifiable errors.

Is it always better to choose a nondiscretionary, mechanical system over a discretionary one? Let us look at some of the advantages and disadvantages of this approach.

Benefits of a Nondiscretionary, Mechanical System A nondiscretionary, mechanical system provides a mathematical edge as determined by testing and adjusting. This is the principle behind the casino and an insurance business, both of which profit from many small profitable trades and occasional losses.

Using a nondiscretionary system avoids emotion. This is an advantage because traders often lose money due to emotional decisions. The nondiscretionary system also reduces other trading pitfalls—overtrading, premature action, no action, and constant decision making. Trading with a properly designed mechanical system also prevents large losses and risk of ruin, which most traders have never quantified or understood. In fact, risk control can be one of the most important advantages of a mechanical system.

Trading with a nondiscretionary system also provides certainty, develops confidence, and produces less stress. Anxiety comes from uncertainty and ambiguity. Although a nondiscretionary system cannot predict the future, it can structure how to react to possible outcomes. It gives a list of responses to events beyond one's control.

Pitfalls to a Nondiscretionary, Mechanical System Although there are many benefits to a nondiscretionary, mechanical system, pitfalls also exist. For one, extrapolating will not have the same results as tests; history does not repeat itself precisely. The more a system is optimized or curve-fitted, the less reliable it will be in the future. In fact, in their book *The Ultimate Trading Guide*, Hill, Pruitt, and Hill (2000) suggest that you should generally expect half the profits and twice the drawdowns as shown in tests of past data. Having been tested, the system designer expects results that are often unrealistic. The designer must be careful not to lose confidence when unrealistic expectations are not achieved.

Nondiscretionary systems often will make profits in clumps, especially if it is a trend-following system. The trader then loses small amounts waiting for the next clump and protecting from large losses. In other words, great creativity may have gone into inventing the system, but its operation is boring. In addition, some system designs allow large drawdowns but still eventually produce profits. The emotional problem for the user is the wait for the drawdown to be recovered and meanwhile the possible loss of confidence in the system. A loss of confidence results in fiddling with the rules or giving up just as the system is about to kick in.

Although a good system adjusts to a changing market, it does require periodic updates. This can often be a source of confusion for the designer. Is it time to update an underperforming system because of a changing marketplace? Alternatively, is the lackluster performance period a time for the trader to sit by patiently waiting for the system to kick in? The answers to these questions are not always obvious.

Remember that the system falls apart if it is not followed precisely. This is what the testing was for, and violations of the rules established from the testing negate the value of the system. This requires considerable discipline.

Using a nondiscretionary, mechanical system is not easy—otherwise, everyone would do it. There is a lot of work in coming up with a system, testing it, adjusting it, and trying it correctly and convincingly. The tendency for many people is to "wing it" and see if it works. That method leaves the trader nowhere.

A Complete Trading System

The following is from The Original Turtle Trading Rules by Turtle member Curtis Faith (bigpicture.typepad.com/comments/files/turtlerules.pdf):

Decisions required for successful trading:

- Markets—What to buy or sell
- Position Sizing—How much to buy or sell
- Entries—When to buy or sell
- Stops—When to get out of a losing position
- Exits—When to get out of a winning position
- Tactics—How to buy or sell

How Do I Design a System?

Now you are convinced that you need to design a system for trading. However, how do you do that? Let us look at some of the requirements and steps involved in creating an effective system.

Requirements for Designing a System

What is needed to design a successful system? Before even considering the components of a system, we must begin with something even more basic—designing a workable, profitable system begins with some basic personal attitudes. Some of the characteristics of the necessary mind-set include the following:

- Understand what a discretionary or nondiscretionary system will do—be realistically knowledgeable, and lean toward a nondiscretionary, mechanical system that can be quantified precisely and for which rules are explicit and constant.
- Do not have an opinion of the market. Profits are made from reacting to the market, not by anticipating it. Without a known structure, the markets cannot be predicted. A mechanical system will react, not predict.
- Realize that losses will occur—keep them small and infrequent.
- Realize that profits will not necessarily occur constantly or consistently.
- Realize that your emotions will tug at your mind and encourage changing or fiddling with the system. Such emotions must be controlled.

- Be organized—winging it will not work.
- Develop a plan consistent with one's time available and investment horizon—daily, weekly, monthly, and yearly.
- Test, test, and test again, without curve-fitting. Most systems fail because they have not been tested or have been overfitted.
- Follow the final tested plan without exception—discipline, discipline, discipline. No one is smarter than the computer, regardless of how painful losses may be and how wide spreads between price and stops may affect one's staying power.

Initial Decisions

Once you are committed to the mind-set and discipline of creating a system, you must make certain decisions about the characteristics of your system. The actual fundamental or technical method used as the basis for the system is relatively unimportant. What is important is that whatever is used can be defined precisely. Most fundamental and technical methods, by themselves, have a sketchy record of performance. Performance in the system will depend more on filters, adjustments, and the entry and exit strategies than the method itself. This does not mean that any old method will work. Pick a model (entry and exit method that has some statistical probability of success) that is familiar, sensible, comfortable, and has a decent record. Be sure it is based on facts, not opinion, and then concentrate on the process of developing a system.

Most systems designers argue that the simpler the system, the better. A system can become bogged down with large numbers of conditions and statistically will lose degrees of freedom, requiring more data and more signals to establish its significance. The market has **entropy**, an inherent disorder that changes periodically in unexpected ways. A system with few variables will reflect the patterns in the market with a certain accuracy. As more variables are added to the system, entropy causes the nonpattern variables to increasingly influence the results, causing the system to eventually decrease profitability because it can account only for the patterns but not the internal market changes. Indeed, when testing a system, the added variables should be tested for their effect on the system results, and if the performance declines, those variables should be eliminated even if they appear logical. Some designers such as Richard Dennis argue against simplicity (Collins, 2005), but they have enormous computer power and knowledge behind them. Hill and others argue that even with modern technology and mathematics, the success of systems now is no greater than the classic systems designed with a hand or crank calculator.

First, you must decide what kind of input and tested model is to be used to generate signals. Some investors depend on fundamental information; most traders depend on technical methods. Others use a combination. The important aspect is to have a clear understanding of the system's premises and to know that the rules will be easily quantifiable and precise. Specificity is much easier to use and to test than generality. You must also understand the logic of the system and be sure that it suits your style of trading or investing.

Second, you must decide on which markets to focus. Is the market suitable for the intended system? Are there opportunities for diversification between markets or instruments? How much volatility and liquidity is required, and what specific instruments will be traded?

Third, you must establish the time horizon for the system. For example, most trend-following systems work better over longer periods, but most pattern systems work in hours and days. Does the system intend to scalp trades, swing trade, or long-term invest? In addition, what is the psychologically best-suited time not only for system logic but also for ease of use? Do you have time to spend all day with the system, or can you monitor the system only daily, weekly, or monthly?

Fourth, you must have a risk control plan; otherwise, you will not know what to do when markets change. Understand that losses are inevitable, but be sure to keep them under control. Admitting losses separates the professional from the amateur. Rationalizing or excusing losses never helps. The market is never wrong—get out, the quicker the better. To do this, devise a stop-loss strategy—"no clinging to the mast of the sinking ship." This strategy should include protective and trailing stops, price targets, and adjustments for volatility, type of market, and any other state that the market might be in. Another option is to have a filter that shuts down the system when the market enters a trading range or has other characteristics that detract from the model's performance. Otherwise, the account may suffer a larger loss. Emotions and judgment become adversely affected, causing missed opportunities, selling profitable positions to get even, and other mistakes. Stop-losses free up nonproductive capital and cause less stress once accepted. In addition to risk control, you must decide whether you should use leverage or pyramiding.

Fifth, establish a time routine, which should include when to update the system and necessary charts, plan new trades, and update exit points for existing trades. As part of your system administration, maintain a trader's notebook, a trader's diary, and a daily equity chart. Maintain a daily trading sheet (similar to an accounting ledger) and a position sheet.

Types of Technical Systems

Technical analysts use a number of types of technical trading systems. Although there are numerous systems, they can be divided into four main categories: trend following, pattern recognition, range trading, and exogenous signals systems.

Trend Following From our knowledge of technical systems, we understand that markets trend at times and trade in a range at other times. The most profitable background is a trending background because the moves are larger and generate fewer transaction costs. While periodically trend trading becomes difficult and many traders begin to believe it is dead, it is not. As Bill Eckhardt, partner of Richard Dennis and originator of the Turtles, has been quoted, "I have lived through the death of trend-following a half dozen times, and, like Mark Twain's death, it was highly exaggerated" (Collins, 2005). Most large-scale mechanical system hedge funds and commodity trading advisors use trend-following systems. Rather than attempting to catch the peaks and valleys, the trend-following system acts in the direction of the trend as soon after it has begun as can be reliably detected. Contrary to the buy low and sell high philosophy, the trend-following system will buy high and sell higher. Schwager believes that slower, longer trend systems work better because the gains are larger, although less frequent, and the whipsaws are minimal. Most trend-following systems add a trend indicator such as the ADX to their set of rules to be sure that a trend is in existence. As we know from earlier studies on trends, the performance of a trend-following system can suffer during a trading range market.

Moving Average Systems The classic trend-following system is composed of two moving averages that generate signals when they cross over each other. In his book *The Definitive Guide to Futures Trading*, Larry Williams discusses how, as early as the 1940s, Donchian demonstrated the validity of this method and showed that it was more successful than the older system of using price versus a single moving average.

If two moving averages are better than one, would three be even better? No, studies have shown that adding more moving averages weakens performance because of the increased number of rules required. Although practitioners frequently report success using moving averages, we must mention that academic studies have shown that moving average crossover systems, even with simple filters, are generally unprofitable. However, academics have not used any kind of risk control in their experiments. Without the use of these important risk-control strategies, the academic studies are not a true measure of the profitability of using a moving average crossover system.

Breakout Systems A variation of the trend-following system is the breakout system. These systems generate buy and sell signals when price moves out of a channel or band. The most popular of these systems is based on a variation of the Donchian channel breakout system or some kind of volatility breakout system using Bollinger Bands or other measures of range volatility. The breakout system can be long term and use weekly figures, or short term, such as the open range breakout systems used intraday.

Problems with Trend-Following Systems Given their profitability, the moving average and breakout strategies are popular. Because many of these trend-following systems are being traded, many others will receive the same signal at roughly the same time and price you will. Liquidity can become strained, and slippage costs from wider spreads and incomplete fills will increase the transaction costs over what may have been anticipated. The solution to this problem is to devise an original system or to spread out or scale entry orders.

Another problem with trend-following systems is that whipsaws are common, especially during a trading range market, as the system attempts to identify the trend. In fact, trend-following systems often produce less than 50% wins because of the many whipsaws during ranging markets. This problem can be reduced with the use of confirmations, such as special price requirements (penetration requirement, time delay, and so on), once a signal has been given, or through filters and diversification into uncorrelated markets.

Inevitably, to avoid whipsaws, a trend-following system will be late in the trend and will thus miss profit potential at both ends of the trend. Unfortunately, this is the cost of a trend-following system. If an attempt is made to clip more profit at each end of the trend, the number of losses will increase from the ranging nature of the trend at its terminal points. On the exit side of a trend, specific trailing stops or such can be used to receive better prices, but again there is the risk of missing another leg in the trend by exiting prematurely.

Losses occur primarily in the trading range preceding the establishment of a trend, as the system tries to identify the next trend as closely as possible. One strategy to combat this is to use a countertrend system at the same time, even if it is not as profitable as the trend-following system. The gains from the countertrend system will offset some of the losses of the trend-following system, and the overall performance results will improve over the trend-following system alone.

Moving-average and breakout systems are usually limited to a one-directional signal only. Part of the advantage in following a trend is to pyramid in the direction of the trend as evidence of its viability becomes stronger. To accomplish this in a trend-following system, other indicators must be used, thus increasing the complexity and decreasing the adaptability of the system.

The greatest fault with trend-following systems is the large percentage of consecutive small losses that produce significant drawdowns. For example, let us say that the system suffers ten small losses in a row while in a trading range. The drawdown to the equity of the account accumulates during this period from the peak of the equity to the subsequent cumulative loss. A series of losses that cause a large drawdown affect not only the pocketbook but also the confidence in the system and often lead to further complications. One strategy to lessen a sequence of losses is the strategy mentioned previously of using a countertrend system. Another is to initiate only small positions on a signal until the trend is well established. Yet another is to run another trend-following system parallel that has a longer or shorter period.

Because a trend-following system often is characterized by clumps of large profits from the trend and many small losses from the trading range, extreme volatility occurs in equity. We will look at this later when we study equity curve smoothness, but the most-often-used countermeasure is to diversify into other markets or systems.

As with most mechanical systems, a trend-following system can work well during testing and then bomb in practice. In most cases, this is due to improper testing and adjusting. Some-times the improper testing is due to unrealistic assumption about transactions costs. Unrealistic assumptions including spreads during fast markets, limit days in the futures market, and other possible anomalies may have given false results during the testing stage of the system under consideration. Remember that the popularity of trend-following systems can affect slippage; this fact often is erroneously ignored in the testing phase.

Occasionally, substantial parameter shifts will occur that the adaptive system will not be able to recognize and accommodate. Again, by diversifying by using more than one system or using market character adjustments to volatility, such problems can be reduced.

Pattern Recognition Systems

"Every ship at the bottom of the sea had plenty of charts" is attributed to noted systems trader Jon Najarian (Patel, 1997). Using patterns requires considerable testing and overcoming the problem of defining patterns. Larger patterns do not succumb to easy computer recognition because of their variable nature. System traders such as Larry Williams, Larry Connor, and Linda Raschke use short-term patterns and limit their exposure with specific position stops and price or time targets. Generally, such systems are partially discretionary because they require some interpretation during the trade entry.

Reversion to the Mean

Reversion to the mean systems are based on the buy-low-sell-high philosophy within a trading range and are also called **trading range** systems. This type of system requires a certain amount of volatility between the peaks and valleys of ranges; otherwise, transaction costs, missing limits, and being stopped out on false moves chew up any potential profits. Generally, these systems are discretionary. They profit from **fading** small counter-trend moves or moves within a flat trend and using oscillators such as the stochastic, relative strength index (RSI), the Moving-Average Convergence/Divergence (MACD), or cycles. The largest potential problem in trading with one of these systems is the possibility of a trend developing that creates the risk of unlimited losses. Protective stops are a necessity.

Generally, this type of system does not perform well. A number of publicly available tests—for example, of buying and selling within Bollinger Bands—have been conducted, and invariably the best performance comes from buying and selling on breakouts from the bands rather than trading within them. The major use of countertrend systems is to run coincident with trend-following systems to dampen the series of losses in the trend-following system during a trading range.

Exogenous Signal Systems

Some systems generate signals from outside the market being traded. Intermarket systems, such as gold prices for the bond market, would be an example of an exogenous signal system. Other examples are sentiment such as the VIX for S&P futures, volume, or open interest warnings of activity that trigger price systems or act as confirmation of price systems, or fundamental signals such as monetary policy or consumer prices.

Which System Is Best?

Which type of system is the best? John R. Hill and George Pruitt, whose business is to test all manner of trading systems (www.futurestruth.com), maintain that the best and most reliable systems are trend-following systems. Within trend-following systems, the breakout systems have the best characteristics—specifically the Bollinger Band breakout systems, and the Don–chian, or channel, breakout systems. Closely behind are the moving-average crossover systems.

How Do I Test a System?

Testing a hypothetical system is absolutely necessary, and the testing process can be tedious because so many ideas of how to trade turn out to be unsuccessful. This is the most difficult aspect of designing a system, and unfortunately, because it is so time-consuming and discouraging, many analysts take short-cuts, such as not performing out-of-sample tests, and end up with a system that eventually blows up on them. The process begins with being sure the data being used in the testing is clean and the same as the data that will be used later when the system goes live. The next is to establish the rules for the model being chosen as the basis for the system and **optimize** the variables chosen. These rules include entry and exit signals at first and will have other filters added later depending on the results of the first series of tests. If a **walk-forward** program is not available, a large portion of the data, called **out-of-sample data**, must be kept aside to use later when testing the system for **robustness**. Once a viable system has been adequately optimized, the resulting parameters are then tested against the out-of-sample data to see if the system works with unknown data and was not the result of **curve-fitting** or **data mining**. This is the disheartening part of system design because invariably the out-of-sample test will fail, and the development must return to the beginning. It is at this point that most amateurs give up.

Clean Data

Not surprisingly, for an accurate evaluation of any system, the data must be impeccable. Without the correct data, the system tests are useless. Data should always be the same as what will be used when the system is running in real time. Not only the data but also the data vendor should be the same source as what will be used in practice. Different vendors receive different data feeds. This is especially a problem in short-term systems, where the sequence of trades is important for execution and for pattern analysis.

The amount of data required depends on the period of the system. A general rule of thumb is that the data must be sufficient to provide at least 30 to 50 trades (entry and exit) and cover periods where the market traveled up, down, and sideways. This will ensure that the test has enough history behind it and enough exposure to different market circumstances.

The real-time trader has enough difficulty with "dirty" data on a live feed, and this becomes just as crucial when testing back data. Cleanliness of data is a necessary requirement. Any anomalies or mispriced quotes will have an effect on the system test and will skew the results in an unrealistic manner. Cleaning of data is not an easy task and often must be relegated to the professional data providers.

Special Data Problems for Futures Systems

Although stock data has a few historical adjustments such as dividend payments, splits, offerings, and so on, the futures market has another more serious problem: which contract to test. Most futures contracts have a limited life span that is short enough not to be useful in testing most systems. The difficulty comes from the difference in price between the price at expiration and the price of the nearest contract on that date into which the position would be rolled. Those prices are rarely the same and are difficult to splice into something realistic that can be used for longer-term price analysis. To test a daily system, for example, two years or more of daily data is required at the very least, but no contract exists that runs back for two years. Of course, testing can be done on nearest contract series, but it is limited to the contract length. This is satisfactory if the system trades minute by minute but not for daily signals in a longer-term system.

To rectify this problem, two principal methods of splicing contract prices of different expirations together in a continuous stream have been used. These methods are known as perpetual contracts and continuous contracts. Neither is perfect, but these methods are the ones most commonly used in longer-term price studies.

Perpetual contracts, also called **constant forward contracts**, are interpolations of the prices of the nearest two contracts. Each is weighted based on the proximity to expiration of the nearest contract to the forward date—say, a constant 90 days. As an example, assume that today is early December, only a few days from expiration of the December contract of a commodity future and a little over three months from the expiration of the March contract, the next nearest. The 90-day perpetual would be calculated by proportioning each contract's current price by the distance each is in time from the date 90 days from now. This weighting in early December favors the March contract price, and each day as we approach the December expiration, the December contract receives less weight until expiration when the perpetual is just the March contract price. The following day, however, the March contract price begins to lose weighting as the June contract price begins to increase its weighting. This process gives a smooth but somewhat unrealistic contract price; it eliminates the problem of huge price gaps at rollover points, but you cannot literally trade a constant forward series. As Schwager points out, "the price pattern of a constant-forward series can easily deviate substantially from the pattern exhibited by the actual traded contracts—a highly undesirable feature" (1996, p. 664).

The continuous, or spread-adjusted, contract is more realistic, but it suffers from the fact that at no time is the price of the continuous contract identical to the actual price because it has been adjusted at each expiration or each rollover date. The continuous contract begins at some time in the past with prices of a nearby contract. A rollover date is determined based on the trader's usual rollover date—say, ten days before expiration. Finally, a cumulative adjustment factor is determined. As time goes on and different contracts roll over to the next contract, this spread between contracts is accumulated and the continuous contract price adjusted accordingly. With this method, the continuous prices are exactly what would have been the cost to the trader had the system signals been followed when they occurred. There is no distortion of prices. Price trends and formations occur just as they would have at the time. The only difference is that the actual prices are not those in the continuous contract. Percentage changes, for example, are not accurate. Nevertheless, the method demonstrates exactly what would have happened to a system during the period of the continuous contract, which is precisely what the systems designer wants to know.

As Schwager points out, "a linked futures price series can only accurately reflect either price *levels*, as does the nearest futures, or price *moves* as does continuous futures, but not both…" (1996, p. 669). Students interested in trading futures can refer to the book *Schwager on Futures: Technical Analysis*, to learn more about these techniques.

Testing Methods and Tools

Fortunately, the wheel need not be reinvented when it comes to testing software. Many trading software products include a testing section. Some are reliable; however, some are not. Before purchasing any such software, you should understand the testing methods and resulting reports of the software. Almost all such programs leave out crucial analysis data and may often define terms and formulas differently from others. For example, the term *drawdown* has different meanings, depending on intraday data, closing data, trade close data, and so forth. You must understand the meaning of all terms in any software program to correctly interpret tests performed by it. With this in mind, the systems analyst must establish exactly what information is desired, what evaluation criteria would be useful, and how the results should be presented.

Test Parameter Ranges

The initial test of a system is run to see if the system has any value and, if so, where the problem areas might lie. When the testing program is run, the parameters selected initially should be tested to see if they fall in a range or are independent spikes that might or might not occur in the future. A parameter range, called the **parameter set**, which gives roughly the same results, bolsters confidence in the appropriateness of the parameter value. If, when the parameter value is changed slightly, the performance results deteriorate rapidly, the parameter will not likely work in the future. It is just an aberration. When the results remain the same or similar, the parameter set is said to be stable—obviously a desirable characteristic.

BOX 1.1 DESIGNING A SYSTEM: "HAL" (NAME OF THE COMPUTER IN *2001: A SPACE ODYSSEY)*

Let us look at a simple case study of how to develop a trading system. Suppose we decide that we will trade International Business Machines (IBM), traditionally a less volatile blue chip. We also decide that we will start with an oscillator called the **Commodity Channel Index (CCI)**. The CCI is an oscillator similar to the Stochastic only it includes a volatility component and thus makes it a more realistic indicator of overbought or oversold. The signals will come from the CCI crossing levels determined by the optimization.

Looking at the monthly chart of IBM (see Figure 1.1) from 2005 through mid-2015, we see several periods of upward and downward trends and trading ranges. This is an ideal history to analyze and test because it includes the three possible trends in any market: up, down, and sideways. It also covers a period of more than nine years, roughly 2250 days, enough to give us plenty of signals.

Normally the CCI is contained with +300 and –300 but is not explicitly bounded. The only variables are the length of the moving averages used in its construction and the level of the two signal lines.

The account will assume a capitalization of $30,000, and commissions and slip-page will be 10 cents per share for each entry and exit or 20 cents per share total. The entries will be limited to 100 shares per trade and only one 100-share position allowed. The reason for this model in our exercise is that we know it has worked well over the past two years, and we want to see if changing the parameters can improve its performance.

Created using TradeStation

FIGURE 1.1 International Business Machines Corporation Common Stock Price (Monthly: January 2005–June 2015).

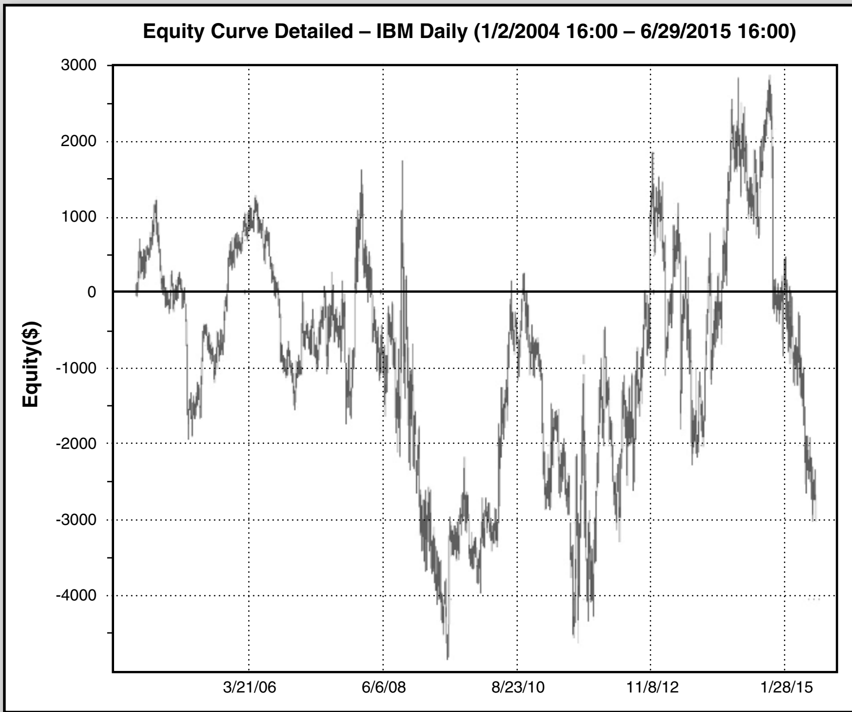

FIGURE 1.2 Equity Curve for Trading IBM Using a Bollinger Band Reversion to the Mean Model (IBM weekly: January 2, 2004–June 29, 2015).

The equity curve for this system with the default length parameter of 14 and signal levels as +100 and –100 is shown in Figure 1.2. An equity curve is a chart of the equity in the account (vertical axis) versus time measured either by trade number or by time (horizontal axis). In Figure 1.2, time is along the horizontal axis. Looking at the chart, we can see that the system had a mixed performance and could easily be discarded as just another oscillator. However, if we change the parameters through optimization and walk-forward testing, perhaps we can find a more reliable formula that worked in the past for the entire period and will have a good chance of working in the future.

The tabulated data in Table 1.1 is from this one run using the standard Bollinger Band parameters.

TABLE 1.1 HAL Initial Test Statistics

Trades	All	Long	Short
Net profit	($2,153)	$2,991	$(5,144)
Gross profit	$32,773	$17,229	$15,544
Gross loss	($34,926)	($14,238)	($20,688)
Profit factor	0.94	1.21	0.75
Number of trades	139	69	70
Percent profitable	57.55	65.22	50.00
Average trade net profit	($15.49)	43.35	($73.49)
Largest winner as a % of gross profit	5.92%	6.32%	12.47%
Largest loser as a % of gross loss	7.76%	19.04%	9.60%

(continued)

TABLE 1.1 ***(Continued)***

Trades	All	Long	Short
Maximum consecutive losing trades	4	3	5
Average days in winning position	14.41	15.2	13.5
Average days in losing position	29.51	26.13	31.83
Buy-and-hold return	80.7%		
Return on account	(40.41 %)		
Monthly average return	($31.63)		
Standard deviation of monthly return	$887.80		
Sharpe ratio	(0.01)		
System MAR (intraday)	(negative)	0.666	(negative)
Trade MAR	(negative)	0.883	(negative)

Source: TradeStation.

Let's look at some of these statistics and learn what they tell us about the HAL system so far:

- **Net profit** is the difference between gross profit and gross loss. It is negative for this system as a whole but positive for long positions. This problem can be attacked in one of two ways: using different parameters for selling short or just using it for long signals. If we use long signals only, we already have a viable system that has worked in the past but not very well. We decide we will adjust both long and short signals with an optimization and walk-forward test.
- **Gross profit and gross loss** are the totals under each category for each trade. Gross profit is the total profit from profitable trades; gross loss is the total from all losing trades.
- The **profit factor** is the absolute value of the ratio of gross profit to gross loss. It shows the profitability of the system. In this case, for every dollar of loss, 0.94 dollars of profit are generated; in other words, it is a losing system. The long side only was favorable at 1.29. The better systems are above 2.00.
- Looking at the **number of trades,** this system generated 139 trades: 69 long trades and 70 short trades. This is a large enough number of trades for reliable statistics. Generally, at least 30–50 trades are required to test a system.
- **Percent profitable** is the percent of all trades that were profitable. In our example, 57.55% of the trades were profitable, yet the system lost money. This suggests that there is something wrong with the losing trades; although fewer in number, they are losing more than the winners.
- **Average trade net profit** is the average profit received per trade. This is negative and suggests that the system is vulnerable to transaction costs.
- The **largest winner or loser versus gross profit or gross loss** figure gives a hint as to whether the gain or loss was accounted for by only one trade. In this case, the largest winning trade accounted for 5.92% of the total gross profit. This is a reasonable size when considering that the total number of trades was greater than 139.
- The **maximum consecutive losing trades** is important because a long string of consecutive losses invariably causes a large drawdown and, thus, a high potential risk for the system. In this case, the number of successive losses is four trades in a row.

It suggests that two whipsaws took place during the test period. Whipsaws can be controlled with stops.

- Considering the **average weeks in winning and losing positions,** there is not much question that the HAL has a problem with losing trades. There should be considerably less time in losing trades. The rule of thumb is that one-quarter of the time can be spent on losing trades versus winning trades, but with a long holding period, the system isn't kicking the losers out soon enough.
- **Buy-and-hold return (80.7%)** is the return gained if the investor bought the IBM on the first day and held it for the entire time period through all its gyrations. This is the number to beat.
- **Return on account (–40.41%)** is the total return on the minimum account size as determined by the maximum drawdown. It should be compared to the buy-and-hold return to see if the system outperforms a do-nothing approach. In this case, the system failed to exceed the do-nothing approach. Of course, such comparisons are not as easy as they look because the concept of risk has not been introduced to either method. The buy-and-hold method has infinite risk because the drawdown can be 100%. The risk of the system has been limited to a much smaller percentage, but we are still observing losses.
- **Average monthly return and standard deviation** of the monthly return are used to determine the volatility of returns. The average monthly return for this system is –$31.63, but it is highly volatile with a standard deviation of $887.80. Ideally, a system should have a standard deviation less than five times the monthly return. In this case, it is 31.5 times, far above the limit, and likely due to the large number of losses.
- The **Sharpe ratio** is a common measure of the return versus risk of a portfolio or system. As we saw in "Selection of Markets and Issues: Trading and Investing," it is a ratio of return—in this case, adjusted for the risk-free return of T-bills, to the standard deviation of return, a proxy for risk. As we stated earlier, however, risk is not just volatility, but is also the risk of capital loss. The Sharpe ratio fails to account for drawdown and fails to account for skewed deviations of return. An investment that deviates more to the upside, for example, will not be fairly represented by the Sharpe ratio, which assumes a normal distribution. These problems are why system designers shy away from the Sharpe ratio and have designed other ratios of return to risk that are more realistic. In this system, the Sharpe ratio is close to zero, suggesting that the return does not exceed the risk-free return of T-bills.
- **System MAR** is the ratio of annual percentage net gain for the system to the maximum percentage drawdown (MDD). The maximum percentage draw-down is the maximum percent that the equity curve corrected from a peak. The ratio measures the greatest decline that occurred during the system run and thus the potential loss in the future for the system. A ratio of greater than 1.0 is preferred.

  Naturally, one wants a system that has no drawdown, but barring that, one wants a system that has profits considerably higher than any drawdown potential. A large drawdown lowers trust in the system and may cause a premature close of the system before it has a chance to perform. The HAL has a negative gain versus a maximum drawdown. The favored standard is anything above 1.00. The ratio is, thus, a gauge for comparing systems.
- **Trade MAR** is the ratio of the net annual gain percentage to the largest trade drawdown in a trade, sometimes called the **Maximum Adverse Excursion.** Where maximum system drawdown may include many trades, the individual trade risk is also needed to gauge the systems performance and isolate where losses are occurring.

HAL may have promise if we can fix the problem with lengthy losing trades and the short selling losses. The long only section is satisfactory and will improve with improvement in the losses, but the entire system is bogged down by poor performance in the short side. This is not surprising in a market that has a generally upward trend, but it still is disappointing. We look next at ways to improve this system with optimized parameters and changes in the model logic.

Optimization

Once you determine that the parameters in your system are valid, you may optimize the system. **Optimizing** is simply changing the parameters of a system to achieve the best results. The most important benefit of optimization is that the designer may find parameters that do not work under any circumstances. If parameters do not work with the past data, it is highly likely they will not work in the future. Thus, optimizing can eliminate useless rules and parameters.

Optimizing is also useful in determining whether certain types of stops are useful. Often the designer finds that there is a limit—for example, to a protective stop—beyond which the stop does not add to the system performance. Often, the distance of trailing stops is too close to the last price, causing premature exits. These determinations can be analyzed more closely with optimization.

Although it can be beneficial, optimization does come with major hazards. With modern computers and sophisticated software, we can take any series of prices and find the best parameters for any predefined system. The problem is that by doing such an optimization, we are just fitting the data to a curve of results and have no idea whether the parameters we have derived will perform in the future. Because the future is what we are attempting to control, most optimization is useless and even dangerous because it gives us a false sense of confidence.

The principal concern with optimization is the tendency to **curve-fit**. Curve-fitting occurs when the optimization program finds the absolute best set of parameters. What the program is really doing is fitting the parameters to the data that is being tested. Thus, it is forming a mathematical model of that data and fitting parameters to that particular time in history. The only way that the parameters will work in the future is if the future exactly duplicates the history that was optimized. Of course, we know this will never happen and, thus, the parameters determined by optimization likely will be useless in the future. Any system could be made to look profitable if optimized; this is a problem that buyers of systems must face when considering purchasing an existing system for investing or trading. The trick is to optimize over a certain period and then test the parameters derived through optimization on a period in which no optimization has been conducted. This is called **out-of-sample (OOS) testing**. Invariably we will find that the results in the optimization will overstate the results in the out-of-sample period and, thus, the optimized parameters should never be used to evaluate the system's usefulness. Optimization should be kept simple. Fine-tuning the system just increases the level of false confidence that eventually will be dashed in real time when the system fails.

There is, thus, some controversy about the use of optimization in arriving at workable mechanical systems. The basic principles of realistic optimization are to keep it simple, test out-of-sample data against in-sample optimization results, preferably use baskets of securities, determine parameter sets instead of single parameters, understand that the best results are high profits with minimal risk, and avoid expecting to find the Holy Grail. Next, we discuss some optimization methods and some tests for statistical significance to perform after the most realistic parameter sets have been determined.

Before optimizing, the analyst must decide what the optimization is looking for in the data. Is it looking for net profit, maximum drawdown, Sharpe ratio, percentage of winning trades, or any other **objective function**? This objective function is an important aspect of the investigation for the best system. What is best? Many analysts use as their objective function a ratio of net profit to maximum drawdown, called the **MAR ratio**, to account not only for profits but also for risk of loss. Others use a **regression line** fit to the resulting profits. A tight fit suggests less volatility and thus less drawdown. A variation is called the **perfect profit correlation**. It assumes that the perfect system would buy every trough and sell every peak and thus generate a certain "perfect" profit. The tested system results are then compared to the perfect system to see how well it correlated.

Methods of Optimizing

As a general rule, an optimization should be done over a considerable period of price data and include those periods when the prices are in trends and in trading ranges. We do not know ahead of time whether the future will be similar, but we do know that there will be trends and trading ranges. Any system must be able to deal with both of these situations and have developed adjustable parameter sets or rules that will account for them. Parameters determined in this manner should be suitable for future conditions.

Whole Sample One method of optimizing is to take the entire price sample and run an optimization of the parameters. This is usually frowned upon because it is the closest to curve-fitting. To avoid curve-fitting, optimization should optimize only a portion of the data, called **in-sample (IS)** data, and test the resulting parameters on another portion of the data, called **out-of-sample (OOS)** data, to see if positive results continue in data not seen before by the optimization process. The selection of data can be a basket of stocks or futures rather than a single market average or issue and should have sufficient data to produce over 30 trades. The diversification of securities reduces the likelihood that any results are solely the result of peculiarities in a particular security, and the large number of signals increases the statistical significance of the results. After determining the optimal parameter sets—those that are consistent and give decent results (but not necessarily the best results)—the next step is to divide the optimization period into segments and run a test on each using the derived parameter sets. The results from these different periods then can be analyzed for consistency to see if the system generated similar results under all conditions. Things to look for are the amount of drawdowns, the number of signals, the number of consecutive losses, the net profit as a percentage of maximum drawdown, and so on. The actual amount of net profit is less important for each stage than are the determinants of risk and the consistency of results (Ruggiero, 2005). If the results are not consistent, the system has a major problem and should be optimized using other means or discarded.

Out-of-Sample Optimization (OOS) This is a method most often used in **neural network** and **regression** studies. We do not cover these particular methods because they are more useful with other data series. They can be used in market analysis, and some people, such as Lou Mendleson (www.profittaker.com), claim to have successfully been able to correlate different markets using neural network patterns. However, for purposes of this study of optimization, we ignore neural networks, multiple regressions, and others such as expert systems and artificial intelligence. Instead, we focus on the most common and productive methods—those used by the majority of systems designers.

One variation of OOS that is commonly used is to take the entire price data series to be optimized and divide it into sections, one of 70%–80% being the IS data and the remaining 20% to 30% being the OOS data. The out-of-sample data can include the first small portion of the total period and the last, or just the last, most recent data. As with all other test methods, the sample must include bull, bear, and consolidation periods. The total amount of data necessary is large in all optimization processes to account for periods of upward, downward, and sideways trends. All must be included so that the system can learn to adjust to any future change in direction or habit.

This method optimizes the in-sample data and then tests it on the out-of-sample data. The out-of-sample results are theoretically what the system should expect in real time. Invariably, the out-of-sample performance will be considerably less than the performance generated in the optimization. If the out-of-sample results are unsatisfactory, the method can be repeated with different parameters, but the more that the out-of-sample results are used as the determinant of parameter sets, the more that the objectivity of the optimization is compromised and the closer to curve-fitting the process becomes. Eventually, if continued in this manner, the out-of-sample data becomes the same as the sample data, and the optimization is just curve-fitting. One other method of reducing the effect of curve-fitting is to use more than one market as the out-of-sample test. It is difficult to have the same parameter set in different markets and at the same time curve-fit. This appears counterintuitive because most analysts would think that each market is different, has its own personality, and requires different parameters. Indeed, when looking at publicly available systems for sale, one method of eliminating a system from consideration is if it has different parameters for different markets. This usually indicates that the results are from curve-fitting, not real-time performance. A reliable system should work in most markets.

Walk Forward Optimization Walk forward optimization is also an OOS method that uses roughly the same price data series as the one described previously. Although there are many variations of this method, the most common procedure is to optimize a small portion of the data and then test it on a small period of subsequent data—for example, daily data over a year is optimized and then tested on the following six months' data. The resulting parameters of this test are recorded, and another year's data is optimized—this time, the in-sample data used includes the earlier OOS data plus six months of the earlier IS data. Again, the results are recorded, and the window is moved forward another six months until the test reaches the most recent data. Each optimization, thus, has an out-of-sample test. The results from all the recordings are then analyzed for consistency, profit, and risk. If some parameter set during the walk forward process suddenly changes, the system is unlikely to work in the future. The final decision about parameter sets is determined from the list of test results.

Optimization and Screening for Parameters We look next at all the different summaries and ratios that a system designer considers in measuring **robustness** (the ability of the system to adjust to changing circumstances), but first we must mention those that are used to screen out the better systems during optimization.

When optimization is conducted on a price series, the results will show a number of different parameter sets and a number of results from each parameter set. We can look at the net profit, the maximum drawdown, and any of the other statistics shown in Box 1.1. Many analysts screen for net profit, return on account, or profit factor as a beginning. They look at the average net profit per trade to see if the system generates trades that will not be adversely affected by transaction costs. Most important, they look at the net profit as a percentage of the maximum drawdown. The means of profiting from a system—any system of investing—are determined by the amount of risk involved. Remember the law of percentages. Risk of capital loss is the most important determinant in profiting. The net profit percentage of maximum drawdown describes quickly the bottom-line performance of the system. Unfortunately, the optimizing software of some commercial systems fails to include this factor, and it must be calculated from other reported statistics.

Measuring System Results for Robustness

When analyzing a system, we look at the system components, the profit, the risk, and the smoothness of the equity curve. We want to know how robust our results are. Robustness simply means how strong and healthy our results are; it refers to how well our results will hold up to changing market conditions. It is important that our system continues to perform well when the market changes because, although markets trend and patterns tend to repeat, the future market conditions will not exactly match the past market conditions that were the basis for our system design.

Components The most important aspect of the optimization and testing process is to be sure that all calculations are correct. This sounds simple, but it is surprising how often this is overlooked and computer program errors have led to improper calculations. The next aspect is to be sure that the number of trades is large enough to make the results significant. The rule of thumb is between 30 and 50 trades in the OOS data, with 50 or more being the ideal. We have mentioned previously that the comparisons between in-sample and out-of-sample results should differ in performance but should not materially differ in average duration of trades, maximum consecutive winners and losers, the worst losing trade, and the average losing trade. We should also be aware of the average trade result in dollars and the parameter stability. We could apply a Student's t-test to the parameters and their results to see if their differences are statistically significant, and we should test for brittleness, the phenomenon when one or more of the rules are never triggered. Once we are satisfied that the preceding inspection shows no material problems, we can look at the performance statistics more closely.

Profit Measures Remember that the point of practicing technical analysis is to make money—or profit. On the surface, it seems as if this is a simple concept: if I end up with more money than I began with, then the system is profitable. Actually, measuring and comparing the profitability of various potential systems is not quite so straightforward. There are several ways in which analysts will measure the profitability of systems. The major ways are as follows:

- Total profit to total loss, called the **profit factor**, is the most commonly used statistic to initially screen for systems from optimization. It must be above 1.0, or the system is losing, and preferably above 2.0. Although a high number suggests greater profits, we must be wary of overly high numbers; generally, a profit factor greater than ten is a warning that the system has been curve-fitted. As a measure of general performance, the profit factor only includes profits and losses, not drawdowns. It, therefore, does not represent statistics on risk.
- **Outlier-adjusted profit** to loss is a profit factor that has been adjusted for the largest profit. Sometimes a system will generate a large profit or loss that is an anomaly. If the profit factor is reduced by this anomaly and ends up below 1.0, the system is a bust because it depended solely on the one large profit. The largest winning trade should not exceed 40% to 50% of total profit.
- **Percentage winning trades** is a number we use on the makeup of risk of ruin. Obviously, the more winning trades there are, the less chance of a run of losses against a position. In trend-following systems, this percentage is often only 30% to 50%. Most systems should look for a winning trade percentage greater than 60%. Any percentage greater than 70% is suspect.

- **Annualized rate of return** is used for relating the results of a system against a market benchmark.
- The **payoff ratio** is a calculation that is also used in the risk of ruin estimate. It is a ratio of the average winning trade to average losing trade. For trend-following systems, it should be greater than 2.0.
- The **length of the average winning trade** to average losing trade should be greater than 1. Otherwise, the system is holding losers too long and not maximizing the use of capital. Greater than 5 is preferable for trend-following systems.
- The **efficiency factor** is the net profit divided by the gross profit (Sepiashvili, 2005). It is a combination of win/loss ratio and wins probability. Successful systems usually are in the range of 38% to 69%—the higher the better. This factor is mostly influenced by the win percentage. It suggests that reducing the number of losing trades is more effective for overall performance than reducing the size of the losses, as through stop-loss orders.

For a system to be robust, we should not see a sudden dip in profit measures when parameters are changed slightly. Stability of results is more important than total profits.

Risk Measures What happens if you find a system that has extraordinarily high profit measures? Chances are you have a system with a lot of risk. Remember, high profits are good, but we must balance them against any increased risk. Some of the major ways that analysts will measure the risk within their system are as follows:

- The **maximum cumulative drawdown** of losing trades can also be thought of as the largest single trade paper loss in a system. The maximum loss from an equity peak is the **maximum drawdown (MDD)**. The rule of thumb is that a maximum drawdown of two times that found in optimizing should be expected and used in anticipated risk calculations.
- The **MAR ratio** is the net profit percent as a ratio to maximum drawdown percent. It is also called the **Recovery Ratio**, and it is one of the best methods of initially screening results from optimization. In any system, the ratio should be above 1.0.
- **Maximum consecutive losses** often affect the maximum drawdown. When this number is large, it suggests multiple losses in the future. It is imperative to find out what occurred in the price history to produce this number if it is large.
- **Large losses** due to price shocks show how the system reacts to price shocks.
- The **longest flat time** demonstrates when money is not in use. It is favorable in that it frees capital for other purposes.
- The **time to recovery** from large drawdowns is a measure of how long it takes to recuperate losses. Ideally, this time should be short and losses recuperated quickly.

- **Maximum favorable and adverse excursions** from list of trades informs the system's designer of how much dispersion exists in trades. It can be used to measure the smoothness of the equity curve but also give hints as to where and how often losing trades occur. Its primary use is to give hints as to where trailing stops should be placed to take advantage of favorable excursions and reduce adverse excursions.
- The popular **Sharpe ratio,** the ratio of excess return (portfolio return minus the T-bill rate of return) divided by the standard deviation of the excess return. The excess rate of return has severe problems when applied to trading systems. First, it does not include the actual annual return but only the average monthly return. Thus, irregularities in the return are not recognized. Second, it does not distinguish between upside and downside fluctuations. As a result, it penalizes upside fluctuations as much as downside fluctuations. Finally, it does not distinguish between intermittent and consecutive losses. A system with a dangerous tendency toward high drawdowns from consecutive losses would not be awarded as high a risk profile as others with intermittent losses of little consequence.

Individual analysts will choose, and even create, the measure of risk that is most important to their trading objectives. Some of the other measures of risk mentioned in the literature are as follows:

- **Return Retracement ratio**—This is the average annualized compounded return divided by MR (maximum of either decline from prior equity peak [that is, worst loss from buying at peak] or worst loss at low point from any time prior).
- **Sterling ratio** (over three years)—This is the arithmetic average of annual net profit divided by average annual maximum drawdown; it is similar to the gain-to-pain ratio.
- **Maximum loss**—This is the worst possible loss from the highest point; using this measure by itself is not recommended because it represents a singular event.
- **Sortino ratio**—This is similar to the Sharpe ratio, but it considers only downside volatility. It is calculated as the ratio of the monthly expected return minus the risk-free rate to the standard deviation of negative returns. It is more realistic than the Sharpe ratio.

Smoothness and the Equity Curve Some analysts prefer to analyze risk in a graphic, visual manner. Two graphs commonly are used as a visual analysis of a system's performance: the **equity curve** and the **underwater curve**.

An equity curve chart is shown in Figure 1.2. It shows the level of equity profit in an account over time. Ideally, the line of the equity profits should be straight and run from a low level at the lower-left corner to a high level at the upper-right corner. Dips in the line are losses either taken or created by drawdowns.

The common measure of smoothness is the standard error of equity values about the linear regression trend drawn through those equity values. Smoothness of a system is affected by changes in the entry parameters or adjustments, such as filters. Because the majority of price action has occurred by the exit, the exit parameters and stops have little effect on smoothness.

The second type of graph used to look at system performance is the underwater curve chart. An example of this type of chart is shown in Figure 1.3. This displays the drawdown from each successively higher peak in equity. It is calculated in percentages and gives a representation not only of how much drawdown occurred, but also of how much time passed until equity recovered from that drawdown. As Figure 1.3 shows, the maximum percentage drawdown in the initial HAL system was a little over 50% of the original capital of $30,000. This chart helps us see that a major problem with the system is not only the size of the drawdowns but also the time it takes for the system to recover. In Box 1.2, we outline a method for improving the system.

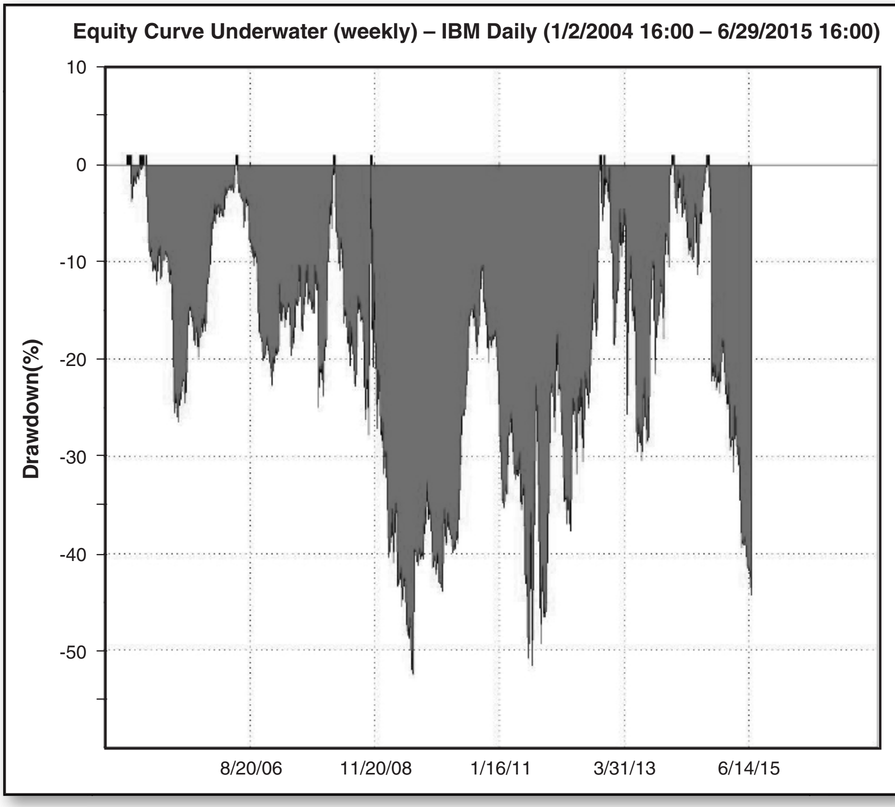

FIGURE 1.3 Weekly Underwater Curve for HAL in Box 1.1.

BOX 1.2 UPGRADE IN THE HAL

Now it is time to upgrade our system based on the results of our initial testing. We first optimize the parameters of the given variables to see if there is a possibility of an improved system just by changing the parameters. This is the first step, and it showed that with curve-fitting, the net profit of over $35,000 was possible versus the loss incurred without adjustments. The second step is that we run a walk-forward test of the results and arrive at a system that we can expect to work in the near future. This is the one we report on here.

The changes made to HAL are threefold. First, we include a filter that will prevent the system from trading when the market is dull. We do this using a requirement that the ADX be higher than its predecessor some unknown number of days prior.

We use the ADX because it is a measure of trend and we don't want to play if there is no trend. There are other configurations of the ADX as a filter, but this is the one that worked best with HAL. Second, we add a percentage protective stop to lower the number of losses that accumulated time and loss while waiting for a buy signal. Third, we run optimizations on the parameters of ADX length, ADX lookback, CCI length, and upper and lower signal levels. The optimal results, using the perfect profit correlation as the objective function, were then run through a walk-forward optimizer to see which combination of parameters has the most likely chance of profiting in the next year.

TABLE 1.2 Tabulated Data for the Final Optimized System for HAL

Trades	All	Long	Short
Net profit	$31,437	$20,294	$11,143
Gross profit	$40,879	$23,179	$17,700
Gross loss	($9,442)	($2,885)	($6,557)
Profit factor	4.33	8.03	2.70
Number of trades	68	32	36
Percent profitable	64.71%	78.13%	52.78
Average trade net profit	$462.31	$634.19	$609.53
Largest winner as a % of gross profit	8.03%	14.17%	16.77%
Largest loser as a % of gross loss	10.19%	31.85%	14.67%
Maximum consecutive losing trades	3	2	5
Average weeks in winning position	45.59	50.44	39.21
Average weeks in losing position	26.54	53.86	15.29
Buy-and-hold return	85.66%		
Return on account	1057.42%		
Monthly average return	$475.07		
Standard deviation of monthly return	$906.07		
Sharpe ratio	0.23		
System MAR (intraday)	5.09	3.68	2.75
Trade MAR (intraday)	10.57	22.08	2.94

Source: TradeStation.

Look at how the system improves with the additions. Figure 1.4 shows the new equity curve for the system. Notice how smooth the curve is now. Net profit has increased from $2,000 to $31,000. The number of trades has decreased because of the ADX filter; the

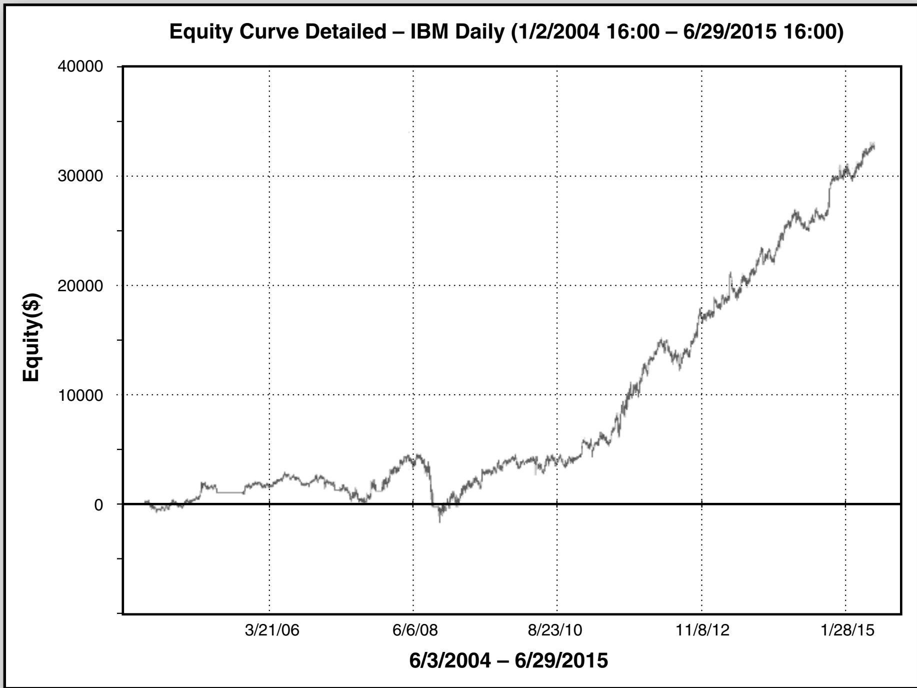

FIGURE 1.4 Equity Curve of Final Optimized HAL.

eliminated trades were obviously losers because the percentage of winners increased. The profit per trade is now high enough to withstand any extra trading costs, and the profit factor is now above the 2.00 standard threshold for a favorable system. The higher monthly return versus the standard deviation is well below the 5.00 normal ratio and explains why the equity curve is so smooth.

Do not use this system as it stands in any stock. It is presented only as an example of the process of looking for parameters, variables, and rules in a system development.

However, we hope that you can see the process of developing a reliable and profitable system and some of the types of adjustments that can be applied to systems—especially the use of stops—to improve performance and reduce risk. System development is a difficult and time-consuming task.

BOX 1.3 WHAT IS A GOOD TRADING SYSTEM?

In his book *Beyond Technical Analysis*, Tushar Chande discusses the characteristics of a good trading system. Chande's Cardinal Rules for a good trading system are the following:

- **Positive expectation**—Greater than 13% annually.
- **Small number of robust trading rules**—Less than ten each is best for entry and exit rules.
- **Able to trade multiple markets**—Can use baskets for determining parameters, but rules should work across similar markets, different stocks, different commodities futures, and so on.
- **Incorporates good risk control**—Minimum risk as defined by drawdown should not be more than 20% and should not last more than nine months.
- **Fully mechanical**—No second-guessing during operation of the system.

Conclusion

Throughout this book, we have looked at a number of technical indicators to guide our buying and selling of securities in reaction to particular market conditions. In this chapter, we turned our attention to mechanizing these reactions. A model is simply a plan or set of rules of when to buy and sell securities. A system uses the model as its base and lets us determine *a priori* how we will react to particular market situations. Having a system in place helps us follow a well-thought-out plan and prevents us from haphazardly trading based on emotion.

Of course, our basic objective in creating a system is to make a profit. Although this sounds like a straightforward goal, the goal of making a profit is not as simplistic as it sounds. Of course, we test our system to see how well it performs. But—and this is an important but—just because a system performs well using past, historical data in a trial situation does not guarantee that we will have the same stellar results in future, real-time trading. The most basic reason for this performance differential is that the market never repeats itself exactly; the system is operating in a different market environment than the one in which it was tested. There are also some system design and testing issues of concern. The system designer must be careful about data choice and not to overfit the data in the sample period. As we have seen in this chapter, even a system that has a high net profit in a test period is not necessarily a system that will perform well in the future. The system designer must consider a host of statistics about the system performance to determine whether the system is suitable for future trading. By following the guidelines laid out in this chapter, you should be ready to design systems and test them to determine their appropriateness for your trading situation.

References

Burnham, Terry. *Mean Markets and Lizard Brains*. New York, NY: John Wiley & Sons, Inc., 2005.
Chande, Tushar. *Beyond Technical Analysis*. New York, NY: John Wiley & Sons, Inc., 2001.
Collins, Art. "The Legend and the Lore of Richard Dennis." *Technical Analysis of Stocks & Commodities* 23, no. 4 (April 2005): 46–53.
Etzkorn, Mark. "Bruce Babcock: Market Realities." *Futures Magazine* (December 1996): 36–42.
Faith, Curtis. *Way of the Turtle: The Secret Methods that Turned Ordinary People into Legendary Traders*. New York, NY: McGraw-Hill, 2007.
Hill, John R., George Pruitt, and Lundy Hill. *The Ultimate Trading Guide*. New York, NY: John Wiley & Sons, Inc., 2000.
Patel, Alpesh. *The Mind of the Trader*. London, UK: Financial Times/Pitman, 1997.
Ruggiero, Murray A., Jr. "Walking Before You Run with System Optimization." *Futures Magazine*. February 2005.
Ruggiero, Murray A., Jr. "Out of Sample, Out of Mind." *Futures Magazine*. April 2005.
Schwager, Jack D. *Schwager on Futures: Technical Analysis*. New York, NY: John Wiley & Sons, Inc., 1996.
Schwager, Jack D. *Technical Analysis*. New York, NY: John Wiley & Sons, Inc., 1996.
Sepiashvili, David, PhD. "How to Best Evaluate System Performance." *Futures Magazine*, March 2005.
Williams, Larry. *The Definitive Guide to Futures Trading*. Brightwaters, NY: Windsor Books, 1988.