Chapter 1

Introduction

This book is about trading using quantitative techniques together with technical analysis. The techniques apply to any of the commonly traded financial issues—stocks, bonds, mutual funds, exchange traded funds, futures, currencies, FOREX, commodities. They are based on analysis of the price and volume of previous transactions made in open markets.

As the subtitle says, this book describes an integrated approach to trading system development and trading management.

As every engineer will tell you, in order to design and develop a product, you must know how it will be used; in order to design and develop a process, you must know how it will be measured and managed.

Our product will be a profitable trading system. Our process will be designing and verifying the system, then monitoring its performance and determining the maximum safe position size. Our metrics will be account growth, normalized for risk.

The purpose of this book is to outline a few relatively simple, but not necessarily simplistic, ideas that will assist readers in their system trading. This book is not intended to be encyclopedic. Our attention will be focused toward developing and managing systems that provide a good trade-off between reward and risk.

Chapter 1 begins with the statement of a goal, briefly discusses a number of topics that provide some background for the development and trading of quantitative systems, and concludes with some reasons why this is so hard.
The Goal

The goal is for the trader to have confidence that the signals generated by the trading system precede trades that provide rewards adequate to compensate for the risk.

The key word is confidence.

The primary limitation is risk.

Major Changes

Some major changes are taking place in trading system development and trading management. This book discusses each, hopefully helping readers understand and prepare.

Broadly—Galileo to Hubble
Charts to Equations
Subjective to Objective
TSDP to Machine learning
Indicators to Patterns
Stationary to Dynamic
Position size into Trading mgmt
Single backtest to Monte Carlo

Frequentist to Bayesian
Idea driven to Data driven
Profit oriented to Risk oriented
Deterministic to Probabilistic
Reaction to Prediction
Decision tree to Non-linear
p-value to Confusion matrix
Equity curve to Distribution

The Process

This is a classical example of fitting a model to a set of data, intending to use the model for prediction.

In order to have a system that generates signals that we have confidence in, that is profitable, and that has acceptable risk, we need several things:

- Data series that have enough variation so that buying and selling produces profit in excess of risk-free alternative uses of the money.
- Those same data series must not have so much volatility that the resulting system has risk of drawdown that exceeds the trader’s personal risk tolerance.
- Existence of patterns in the data that precede profitable trading opportunities.
- Existence of a set of rules and parameters, call it a model, that recognizes the patterns and issues trading signals.
- Our ability to discover the model and verify that it works for historical data, creating a trading system.
Our ability to monitor the performance of the trades generated by the system over time to verify that the system has learned to recognize the patterns and that the patterns continue.

Our ability to compute the correct position size for each trade so that we maximize account growth while holding drawdown within personal limits of risk tolerance.

Our ability to recognize system breakdown and take the system offline before excessive loss of trading capital.

One paragraph to describe the goal. 2000 pages and counting (this is the fifth book in the series) to describe the process. We begin with some review.

**Premises of Technical Analysis**

The underlying assumptions of technical analysis are:

- The markets are not completely efficient.
- There is information—patterns—in the historical price series that can be used to identify profitable trading opportunities.
- Trading systems can be designed to recognize the patterns and give buy and sell signals.
- Patterns similar to those found in the historical data will continue to be found in future data.

**Two Components of Trading**

I assume that readers want to actually trade—to buy and sell some financial assets. The money is made or lost by trading, based on signals to buy and to sell that come from the system. There are two distinct components:

- Developing the system that generates the signals.
- Managing the business of buying and selling.

The flowchart in Figure 1.1 illustrates those two components, the subtasks, the sequence in which they are performed, and the interrelationships.
Development handles issue and data selection; and the design, testing, and validation of the trading model. That includes calculation of indicators, establishment of rules for trade entry and trade exit, searching
to detect patterns, metrics for measuring success, testing to validate the pattern recognition, and establishing a baseline with which to compare future performance.

Trading management focuses on monitoring the health of the system being traded, estimating risk, determining position size, estimating profit potential, and making the trades. When performance begins to decline, it may be necessary to return it to development.

The two components share a common element—the set of trades that, during development, is the best estimate of future performance, and, during trading, is that best estimate set of trades augmented by trades actually taken.

Each side of the flowchart has its own model.

Trading system development has long been thought of as a relatively simple process of applying some chart pattern or indicator to a lengthy series of historical price data, often including a search for the best rules and parameter values, and often including calculation of the position size to be used for each trade.

As both the thinking about that process and the tools available for use with that process have evolved, we can develop better trading systems and better trading management systems by separating the single system into two distinct components.

The first is the trading system. It models the left column—the one labeled as development. Its input is one or more series of prices and its output is a series of buy and sell signals and the resulting trades. It is discussed in the next few pages of this chapter, and in detail in Chapters 6, 7, and 8.

The second is the trading management system. It models the right column—the one labeled trading. Its input is a set of trades and its output is the position size for the next trade. It is discussed a little later in this chapter, and in detail in Chapter 9.

**Position sizing**

First a few words about position sizing—determining the number of shares or contracts for the next trade.

Position sizing is vitally important. Risk of account-destroying drawdown and opportunity for account-growing profit are closely linked, with position size the critical and coordinating variable.

Including position size while evaluating the data series prior to trading model development (as explained in Chapter 5, Issue Selection), and again in trading management (as explained in Chapter 9, Trading Management), is correct and important. Those two chapters discuss techniques in detail.
But position size should not be a component of the trading model itself (any of those models described in Chapters 6, 7, and 8). Including it there causes two problems:

1. During development of the trading system, using position size other than a fixed number of contracts or dollars introduces a bias that favors specific models (rules and parameter values) that benefit from specific order of trades that happen to occur in the historical data used for development but are unlikely to be repeated in the future, as compared with models whose trades are evaluated independent of order.

2. The trading system has a large number of rules and parameters that can be varied in the search for the best system. The management system has only one—position size. Including position size in the trading system removes it from the management system, leaving no variable that can be used for objective trading management.

Part of the problem of where to put position size calculation was, in the past, due to limitations of available trading system development platforms.

**Position size calculations depend on estimation of distributions of results.**

Trading system development platforms are very good at processing price data, producing a trade sequence, and computing single-valued metrics. But relatively poor at the complex mathematics and Monte Carlo simulations required to estimate and analyze distributions related to the trades. Until recently, there was little choice—include position size in the trading system or deal with it using a separate process and a separate analysis program—perhaps a spreadsheet.

Recent advances in software have provided new opportunities. General purpose languages, such as Python, have been augmented with libraries such as Pandas to ease the handling of time series data, and with libraries such as NumPy and SciPy to ease complex mathematics and Monte Carlo simulations, and with libraries such as Scikit-learn to assist in pattern recognition and classification. As we will see in Chapter 8, Python can be used as a trading system development platform.

The technological barriers to more accurately modeling trading systems and trading management are being removed. Dynamic position sizing, which is discussed in Chapter 9, can now be implemented in the same program that processes the price data and generates trading signals. The appearance is that position sizing is being added back into the trading system. In reality, it is that the two separate phases—development and management—can be together in a single program. The trading system recognizes the patterns in the data and issues the buy and sell signals, and the management system determines system health.
and correct position size. They do this through sharing the best estimate set of trades. This is something that was not available to ordinary trading system developers and traders just a few years ago.

## Development

### System = Model + Data

As illustrated in Figure 1.2 a trading system is a combination of a model and some data.

![Schematic of a Trading System](image)

*Figure 1.2 Organization of a trading system*

### The Model

The model accepts the data series that are being used. The data always includes a primary series—the issue being traded. It may also include auxiliary data, such as that used for inter-market analysis.
The model performs whatever data alignment and transformations are necessary. Parameters are chosen and indicators computed. The logic, rules, and parameters define patterns. When those patterns are found in the data, entry and exit signals are issued.

**The purpose of the model is to recognize patterns that precede profitable trading opportunities.**

The output from the model is a list of trades for the time period being tested, together with a summary of performance.

The model does not include any position sizing—that is handled in trading management.

Chapter 6 discusses general issues related to development of the model. Chapters 7 and 8 discuss issues related to model development using indicator-based techniques and machine learning based techniques, respectively.

**The Data**

**Primary data**

The primary data series is a time-ordered sequence of quotations. Each quotation represents the price of the issue being traded. The prices can be:

- Individual transaction prices—ticks.
- Individual quotations—bid and ask.
- A set of values that provide the range of prices for some period of time—a bar.

The data format assumed throughout this book is bars. Each bar represents a fixed length of time and is a set of numbers that specify the prices associated with that bar. When the issue is a stock, ETF, or commodity, the prices typically include the first, highest, lowest, and last for that period, referred to as open, high, low, and close. Note that, in general, we cannot assume that the first price occurred at the time the bar opened, nor that the last price occurred at the time the bar closed. We never know, nor should we assume, anything about the order of prices within the bar. Specifically, without examining bars of shorter time duration within a longer bar, we cannot determine whether the high came before the low or after it.

In some cases, such as with stocks and ETFs, the volume of shares is also included and reported.

The most common bar length is one trading day, in which case the data is described as being *daily bars* or *end-of-day* data. Bars can be as short as one second, or a few seconds, one minute, or some number of minutes. Any of these is described as *intra-day* data or intra-day bars.
When buy and sell signals are issued and trades created, the transaction prices come from this primary data series.

Chapter 5 discusses selection of the primary data series.

**Auxiliary data**

In addition to the prices of the issue being traded, the model might use auxiliary data. That could be the price of other tradable issues—for example, including the price of bonds in a model that trades stocks. Or it could be a non-price series—for example, the number of advancing or declining issues from a broader set, such as a market sector.

Before being analyzed, all data series must be aligned by the date and time associated with each bar, and any missing data filled in. The obvious choice to provide the master list of dates and times is the primary data series.

**Assume nothing about distribution**

Numerous studies have documented that financial data does not reliably follow any of the standard statistical distributions. In general, we do not know—or even need to know—the distribution of the data. It is important to accept the data as it is without making additional assumptions as to being normal, log-normal, or any other distribution.

**Distributions**

Emanuel Derman writes: “Models are simplifications, and simplifications can be dangerous.”\(^1\) The point I hope to make in this section is that systems developers should avoid simplification of data representation. In short—whenever possible use distributions rather than a limited number of scalar values.

The information content that describes a trading system over a given period of time can be described in many ways. The following list is in decreasing order of information.

- Reality. Trades, in sequence, that actually result from applying the system.
- List of trades, in time sequence.
- Set of trades.
- Distribution of trades.
- Four moments describing the distribution.
- Mean and standard deviation.
- Mean.
- Direction.

Probability and statistics distinguish between population and sample. The population is all items of the type being analyzed. The sample is a subset of the population that has been observed. The purpose of developing trading systems is to learn as much as possible about the population of trades that will occur in the future and make estimates of future performance. The results of testing trading systems form the sample that is used to make those estimates.

**Reality**

Reality cannot be known in advance. Estimating reality, the population, is the purpose of system validation. Reality is the logic of the system processing the future data series.

**List of trades, in time sequence**

The list of trades, in time sequence, that results from processing a data series that is similar to the future data, is the best estimate we can obtain of reality. There is one of these sequences for each unique series of test data and each set of logic and parameter values. Using these results to estimate future profitability and risk depends on the degree of similarity between the test data and the future data.

**Set of trades**

The set of trades, ignoring time sequence, relaxes the assumption of the trades occurring in a particular sequence. It provides a set of trade data with, hopefully, the same characteristics as the future data, such as amount won or lost per trade, holding period, intra-trade drawdown, and frequency of trading. Selecting trades from this set in random order gives an opportunity to evaluate the effects of similar conditions, but in different time sequence.

**Distribution of trades**

A distribution can be formed using any of the metrics of the individual trades. The distribution is a further simplification since there are fewer (or at most the same number of) categories for the distribution than for the data used to form it. For example, a distribution of percentage gain per trade is formed by sorting the individual trades according to gain per trade, establishing ranges and bins, assigning each trade to a bin, and counting the number of trades in each bin. A plot of the count per bin versus gain per bin gives a plot of the probability mass function (often called the probability density function, pdf).

**Four moments**

Distributions can be described by their moments. The four moments most commonly used are named mean, variance, skewness, and kur-
tosis. Depending on the distribution, some or all of the moments may be undefined.

- **Mean.** The first moment. The arithmetic average of the data points.
- **Variance.** Second moment. A measure of the deviation of data points from the mean. Standard deviation is the positive square root of variance.
- **Skewness.** Third moment. A measure of the lopsidedness of the distribution.
- **Kurtosis.** Fourth moment. A measure of the peakedness and tail weight of the distribution.

**Mean and standard deviation**

Mean and standard deviation are commonly computed and used to describe trade results. They can be used in the definition of metrics such as Bollinger bands, z-score, Sharpe ratio, mean-variance portfolio, etc.

**Mean**

The mean gives the average of the values. Mean can be computed in several ways, such as arithmetic mean and geometric mean. Median is an alternative measure of central tendency of a sample that is often useful.

**Direction**

Direction of a trade describes whether it was a winning trade or a losing trade. Direction is meant to represent any way of describing the trades in a binary fashion. Other ways might be whether the result was large or small in absolute value, or whether the maximum favorable excursion met some criterion, etc.

**Stay High on the List**

With each step down this list, a larger number of data points are consolidated into a smaller number of categories, and information is irretrievably lost. Knowing only the information available at one level makes it impossible to know anything definite about the population that could be determined at a higher level. Working with only the mean tells us nothing about variability. Working with only mean and standard deviation tells us nothing about the heaviness of the tails. Using the four values of the first four moments enables us to calculate some information about the shape of the population, but nothing about the lumpiness or gaps that may exist.
For more discussion and examples, read Sam Savage\(^\text{2}\) or Patrick Leach.\(^\text{3}\)

**Patterns**

We will be examining data looking for patterns, for profitable trades, and the relationship between the patterns and the trades. The patterns will be described as a set of rules and coded into the model.

**Non-stationary**

Stationarity is a feature of data that refers to how a particular metric of the data remains relatively constant or changes as different subsets of the data are analyzed.

A process, or set of data, is described as *strictly* or *strongly* stationary when the distribution and its parameters—such as mean, variance, skew, and kurtosis—do not change over time, including showing no trends. Stationarity is a required assumption for some analysis techniques.

The techniques discussed in this book extend the concept of stationarity to whatever metric is being analyzed. In particular, we will be careful to avoid disturbing the relationship between critical patterns and the trades that follow. We want that relationship to remain stationary.

Traditional statistical analysis, including much of both probability and machine learning, assumes the data being analyzed is strongly stationary. The theorems upon which the techniques are based, in particular those that give limits to accuracy and/or error, often require strong stationarity. That assumption is reasonable for applications such as political polling, medical diagnosis, and character recognition.

But time series data is seldom stationary, and financial time series data is never stationary. **Be cautious when applying any technique that assumes the data is stationary to financial time series—there will probably be an undesirable bias.**

**Synchronization**

The model specifies the logic, rules, and parameters. The rules, for example, might be to enter and exit as two moving averages cross. The parameters include the lengths of the two moving averages.

The data is the price history of the issue being traded, perhaps augmented by other data series.

---


A trading system is profitable as long as the logic identifies patterns in the data that precede profitable trading opportunities. That is, as long as the logic and data remain synchronized.

The logic of a typical trading system is relatively fixed. It is designed to detect a particular set of patterns. The data change, following changes in areas that affect the issue—economics, politics, weather, etc.

As the data changes, the patterns in the data move in and out of synchronization with the logic. When synchronized, the system is healthy, it is profitable, gains are steady, drawdowns are low; when unsynchronized, the system is broken, it is unprofitable, gains are sporadic, drawdowns are high. The profit potential and drawdown risk of a system are determined by the accuracy with which the system identifies the patterns.

**During periods of close synchronization, the system is healthy and large positions may safely be taken. As synchronization weakens, position size must be reduced.**

**Signal and Noise**

The data consists of two components—signal and noise. The signal consists of the patterns we hope to identify. What constitutes signal is determined by the model. Everything that a particular model does not explicitly consider to be signal is noise and interferes with identification of the signal patterns.

For a book-length discussion of the relationship between signal and noise in a wide variety of applications, I highly recommend Nate Silver. 4

**Data Series are Not Interchangeable**

It is the combination of a model and some data that comprise a trading system.

Just as we cannot expect different models to be equally effective for a given data series, we cannot expect a given model to be equally effective applied to different data series. If one model does work for a wide range of data, that is a plus. But it is not a requirement.

**Learning**

Learning is the process of examining data, recognizing some patterns, observing related patterns, and hoping there is a generalization. As it applies to trading systems, we will be looking for patterns that provide either:

---

• Classification. Buy or Sell. Either to open a new position or to close an existing position. The issue of how much to buy or sell—position sizing—is addressed separately from pattern recognition.

• Estimation. Direction and magnitude of change anticipated. If an estimation technique is used, we may apply a threshold filter to convert the estimation into a classification category.

Learning is not possible unless there is data to be examined and patterns to find. Preferably a lot of data and a lot of patterns. This is definitely a data mining activity. The data mined is called the in-sample data. We are searching for patterns within the historical price data that precede profitable trades.

We cannot learn a feature that has not been seen. The in-sample data must include examples of the patterns to be learned.

Identifying patterns in the in-sample data is necessary for learning, but not sufficient. There must be generalization. The test for generalization is validation. That is, testing previously unused data to estimate the success of detecting the patterns and defining the rules. The data tested for validation is called the out-of-sample data.

Validation is the step designed to provide the confidence requested in the goal.

Subjectivity and Objective Functions

There are many subjective decisions to be made.

Every day, traders must make decisions:
• Whether the system is healthy.
• When to enter.
• How large a position to take.
• When to exit.

Discretionary traders acknowledge the subjectivity associated with those decisions and draw on experience.

Systematic traders use objective functions designed to identify important decision criteria and quantify them. An objective function is alternatively called a loss function, cost function, utility function, fitness function, or optimization metric. An objective function is a formula that includes terms for each of the criteria or variables important to the decision.

Weights proportional to the importance of the criteria are given to each of the terms, and the terms added together resulting in a single numeric quantity—an objective function score. The score is computed for each alternative being evaluated. The alternatives are sorted according to their score. Providing the objective function has been well designed,
the order of subjective preference is the same as the order of objective function score.

Objective functions are important in both phases of trading:
• In development—to rank alternative systems.
• In trading management—to decide the size of the next position.

**Rank Alternative Systems**

A trading system is a set of computations, logic statements, and parameter values that comprise a set of rules that identify profitable trading patterns and give buy and sell signals.

There are an infinite number of possible systems.

In order to make the process manageable, relatively simple systems are designed to focus on specific trading ideas, such as trend following, mean reverting, seasonality, etc.

For any one of these ideas, there are many alternatives. A trend following system might have logic that looks at breakouts, or the crossing of two moving averages, or the projection of a regression. For each of these there are numeric parameters such as the lengths of the moving averages, or magnitude of breakout. There might be multiple rules to exit a position, such as logic, trailing exit, profit target, and / or maximum loss stop.

Designing a trading system is an iterative process of:
• Modify the logic and parameters.
• Test the performance.

Each set of logic and parameters, together with the data series, creates a new trading system—one of the alternatives to be evaluated.

The developer needs to decide which is best, and best is subjective. The purpose of the objective function is to provide an objective metric that represents the subjectivity of the developer’s definition of best. The objective function she uses includes terms for important features such as gain-per-trade, holding period, and maximum loss.

**Estimate Distributions of Performance**

The trading system that results from the design, testing, and validation provides a single set of trades with single mean, single standard deviation, single terminal wealth, single maximum drawdown.

These results will be repeated as the system is traded only if future prices are exactly the same as the historical series used during development. In order to estimate profit potential and risk it is important to consider the distribution of potential results.

Modeling future performance, including evaluating system health, estimating risk, and estimating profit potential, is based on:
• Using the set of trade results that, in the judgment of the developer, best represents the trades that are likely to occur in the future.
• Using Monte Carlo simulation techniques to create many equally likely trade sequences.
• Analyzing the distributions of drawdown and profit resulting from the trade sequences.
• Comparing both the magnitude and probability of both the drawdown and profit potential with the trader’s personal tolerance for risk and desire for profit to determine system health and position size.

Beginning with determining the maximum safe position size that normalizes the risk associated with a set of trades to keep it within your personal risk tolerance, an objective function based on the Compound Annual Rate of Return (CAR) at some confidence level, say the 25th percentile, is a nearly universal objective function. It is very useful in deciding whether a system is worth trading, and in comparing performance among alternative systems.

The process is outlined in Chapter 2 of this book. For an in-depth explanation of the Monte Carlo method used, including the free software necessary to run the Monte Carlo simulations, see my Modeling book.5

Trader Psychology

We often hear of the importance of psychology in successful trading. The need for the trader to understand himself, to trust the system, to take all signals, to enter the market when the buy signal appears, to set stops at a comfortable level, to exit the trade when the money management stop is hit, to exit the trade at a profit when the profit target is hit, to keep trading through drawdowns. And if the trader considers second-guessing the system, he should consult a trading coach to help him realign his beliefs and accept the system.

In my opinion, that is exactly backwards. We all have personal beliefs about the way the markets work, comfort levels with risk, and preferences related to trading. We know what trading frequency fits in with our other activities. We know what level of drawdown causes us to lose sleep. We can incorporate our own preferences into the objective function we design for our own use when developing and trading our own systems.

A system that scores high marks using our own custom objective function is already one we can expect to be comfortable using. A well designed custom objective function goes a long way toward

5  Bandy, Howard, Modeling Trading System Performance, Blue Owl Press, 2011.
avoiding the cognitive dissonance that requires professional consultation to cure.

Trading management

The trading management sections of this book discuss a new and unique technique, *dynamic position sizing*, and introduce a new metric of system health, *safe-f*.

Position sizing is widely recognized as an important component of trading. The position sizing methods most widely discussed to date make oversimplifying assumptions. They either assume that position size is a stationary variable and a single position size can be applied to a trading system without need for periodic recalculation; or they assume that position size can be determined from within the system's model, then include the position sizing calculation with the logic and rules. Neither is true. Use of either increases the likelihood of serious equity drawdown. Position size is not stationary—position size varies as the health of the system varies. Position size cannot be determined from within the model without outside reference.

Dynamic position sizing monitors system performance trade by trade. Using Monte Carlo simulation and Bayesian analysis, it determines risk of drawdown, assesses the personal risk tolerance of the trader, computes safe-f—the maximum safe position size for the next trade—and estimates profit potential. All on a trade-by-trade basis. *Safe-f* gives you a clear indication of system health, including when the system should be taken offline.

The correct position size for system that is broken is zero.

Chapter 9 discusses trading management.

Risk Tolerance

Everyone has a personal tolerance for risk. Every data series has some inherent risk, independent of the model. Every trading system has some risk. In Chapter 2, we give some techniques for assessing and quantifying personal risk tolerance, for assessing the risk associated with a data series, and for a trading system.

Why Traders Stop Trading

Assume a trader has a method—mechanical, discretionary, or a combination of both—that she has been using successfully. Also assume that she understands both herself and the business of trading, and wants to continue trading. Why would she stop trading that particular system?

Here are a few possibilities:
1. The results are too good.
2. The results are not worth the effort.
3. The results are not worth the stress.
4. She has enough money.
5. There is a serious drawdown

1. Results are too good

She is afraid that this cannot possibly continue.

Her system—any system—works when the logic and the data are synchronized. There are many reasons why systems fail and should be taken offline, but a sequence of winning trades should be seen as a success.

She should continue trading it until one of the other reasons to stop happens.

2. Results are not worth the effort

There is not much gain, but not much loss either. Other things in life are more important. On balance, the time, energy, and resources would be more productively applied doing something else.

3. Results are not worth the stress

Performance is satisfactory, but at a high cost—worry and loss of sleep. Regardless of the position size indicated by the distribution of risk, the positions being taken are too large.

She should either reduce position size or have someone else execute the trades.

4. She has enough money

Not matter how good a system is, there is always a risk of serious loss. When she has reached her goal, she should retire.

5. There is a serious drawdown

The magnitude of the drawdown needed for it to be classified as serious is subjective. Among my colleagues and clients, those who manage other people’s money typically want drawdown limited to single digits. Those trading their own money may be willing to suffer drawdowns of 15 or 20 percent.

But there is a level at which everyone stops trading the system—preferably while the account still has a positive balance.

My view is that experiencing a large drawdown is the primary reason people stop trading a system.

What causes a large drawdown and how should the trader react to it?
• The system is broken.
• There was an unexpected sequence of losing trades.
• The system is out of sync.
• The position size is too high.

As the account balance drops from an equity high into a drawdown, it is not possible to determine which is The reason.

All of the reasons are true to some extent. A system that is broken breaks because the logic and the data become unsynchronized, causing an unexpected sequence of losing trades and at a time when position size was too high for conditions.

The solution is two-fold.
1. Continually monitor system performance and system health.
2. Modify position size to reflect recent performance.

During the trading system development process, a baseline of system performance is established. Using the out-of-sample trades from the walk forward phase is a good source of this data. Personal risk tolerance and system risk, taken together, determine position size for that system performance. As system performance changes, position size must also change.

Position size varies in response to system health.

Do not continue to trade a system that has entered a serious drawdown expecting that it will recover. It may recover on its own; it may require readjustment; or it may be permanently broken and never work again.

Take it offline and either observe it until recent paper-trade results demonstrate that it is healthy again, or send it back to development.

The correct position size for a system that is broken is zero.

Confidence

In the end, you must have confidence. If not confidence, then faith.

The forums that discuss trading systems and their development often ask about the value of walk forward testing. The question is usually accompanied by comments about how hard it is to get good results from the out-of-sample tests from the walk forward runs, whereas it is relatively easy to get good results from optimization and backtesting.

My first reaction is the obvious one—it is hard to get good out-of-sample results because the markets are nearly efficient and it is hard to write a set of rules that detect an inefficiency in advance.

But the first question leads to a deeper consideration about trading systems and trading. Having confidence in a system.
It is my view that the universe of trading system application divides into two—having confidence and having faith.

If you want quantifiable confidence—the kind that tells you whether to hit soft 17 at blackjack, or to hit the blot in your inner table in backgammon, or to buy a recent low, or to buy a new high breakout—my techniques are designed to provide quantifiable confidence in both development and trading.

The problem is harder than it looks at first blush. The characteristics of a trading system determine to a large extent whether it is even possible to have confidence. In order to be useful, there must be enough data points—closed trades or daily account equity values—to compute useful statistical metrics. Examples of useful statements about confidence are:

- To put a low p-value on a set of system results, such as: “we can reject the hypothesis that the expectancy is less than 0.0 with a p-value of 0.05.”
- To put limits on estimates, such as: “with 95% confidence, the worst maximum drawdown for the next year for an account with an initial balance of $100,000 trading at a fraction of 0.80 is 20%.”

Statistical metrics such as these can be computed for any data set—real or hypothetical. If future trades will be made based on these statistics, the data set used to compute the test statistics must be as unbiased as possible.

Using the walk forward technique with trading systems that trade frequently and have short holding periods gives the trading system developer a reasonable chance of producing a set of trades that is both large enough and unbiased enough. Even at that, it is all too easy to introduce bias—bias that will cause reward to be overestimated and risk underestimated—into even the walk forward out-of-sample results.

Compare with backtesting with little or no out-of-sample testing, (which is the all-too-common method in both the popular trading journals and many professional publications), or with systems that have such long holding periods or infrequent trading that an unbiased data set cannot, for practical purposes, be produced.

When in doubt, test it! Do not accept traditional wisdom blindly. Is the 200 day moving average is a good trend indicator? Is trend following the best system to grow a trading account with low drawdown? Those rules may be good ones, and they may lead to trading systems that are appropriate for your use. Or they may not.

Test everything yourself. Your logic, your data, your execution, your estimation of system health. If those tests give you confidence, act accordingly.
Beware of following the advice of the White Queen: “Why, sometimes I’ve believed as many as six impossible things before breakfast.”

If you must act on faith, ask yourself how the casino can build such a fine facility. Stand next to the roulette wheel and listen to the young man tell his partner: “There have been six reds in a row. Black is due.”

**Decisions and Uncertainty**

Most of the decisions we make in life are choices that involve weighing opportunity against risk. Most of the calculations are extremely complex and involve estimating costs and values of things not easily quantified—whom to choose as a partner, where to live, what employment to pursue. All are specific applications of making decisions under uncertain conditions. It seems that the more important the decision, the less opportunity we have to practice and the more important it is to be correct early in the process.

How we handle our finances is certainly an important area, and one where we don’t get many practice runs. For traders, the goal is maximizing trading profits while minimizing the risk of bankruptcy. In the spectrum of life’s activities, this is a problem that is relatively easy to quantify and analyze. The major aspects already have easily measured units of value—dollars. And, given a little understanding of probability and statistics, along with some computer data analysis, we can outline a plan.

**Why This is So Hard**

Developing profitable trading systems is a difficult problem for many reasons.

**Low Signal to Noise Ratio**

The data is very noisy. The markets are nearly efficient. Thinking of the patterns we are searching for as the signal, the signal is weak and is hidden among a lot of noise.

**Nonstationary**

The data is not stationary. Nothing stays the same for long. The characteristics of the data change over time. A solution found for one period of time may no longer apply in a later time period. Determining the appropriate lengths of time to use for the in-sample and out-of-sample periods is difficult.

---

Time Series is Different

Time series data, and particularly financial time series data, is different than the data typically fed to models. The vast majority of modeling, simulation, statistics, and machine learning books and articles assume the data is stationary.

When it is, models that learn and predict accurately are relatively easy to build. The theory provides guidelines, and in some cases rigorous estimates, of out-of-sample performance.

When it is not, techniques that rely on stationarity still give results. But the theoretical justification fails to hold. Results overestimate profit and underestimate risk. Real-time trading results are much poorer than anticipated.

Feedback

The purpose of a trading system is to recognize an inefficiency in price, then make trades that capture that inefficiency. An example is the process described as “arbitraging an inefficiency out of the system.” In the process, price becomes more efficient and more difficult to detect in the future.

Trend Following

Every trade is a trend following trade. No matter how the entry is made, the price must change in the direction predicted in order for the system to be profitable. The trend must complete its expected or predicted run before there is a drawdown or early exit from the trade. As more traders recognize that particular trending pattern, the trend becomes shorter in both time and price change.

Limited Profit Potential

The markets are very nearly efficient. Every successful trade removes some inefficiency and makes future profitability less likely. Given a desirable trend, the first positions taken get the best price. Later fills are at worse prices. The latest trades do not obtain enough profit to cover commission and slippage.

Different Systems, Same Trades

Trades can be categorized according to the amount of change from entry to exit, or the amount of time they are held. Over a period of time, there are only a few profitable trades for any given trade profile. Everyone developing systems that will hold trades for one to five days or one to two percent will locate the same profitable trades no matter what pattern or entry technique they are using.
Very Large Search Space

There are many potential solutions. Patterns can be described in terms of indicators, seasonality, candles, etc. Finding a pattern that works is a search among a large number of possibilities. It is very easy to overfit the model to the data.

The in-sample results are always good. With so many variables available to fit so few data points, it is easy to obtain good in-sample results.

Out-of-sample Results Matter

Out-of-sample results are the important ones. They may not be good for two reasons.
- One. The system was never good. The rules fit the in-sample noise rather than meaningful and predictive patterns.
- Two. The system is no longer good. The characteristics of the data have changed since the patterns were identified.

Financial Rewards

Rewards for success are high.

Barriers to entry are low.

Trading is competitive. Trading is nearly zero sum. My profit is some other trader’s loss. Knowledge shared by one trader reduces his future profit.

Competition

There are no handicaps. Novice and journeyman golfers, bowlers, tennis players, chess players, bicycle racers, and go players can all enter tournaments knowing that they will either be competing with people whose skill level is roughly the same as their own, or they will be given a handicap that compensates for their lack of skill and experience. Not so for traders. When any of us takes a long position in a stock, ETF, or futures contract, the person taking the opposite position is very likely to be trading for a major financial institution. They are well educated, well equipped, well funded, well supported, and are using the best methods and systems.
Figure 1.1  After I win a few races, I am going to buy a really good bicycle and enter the Tour de France.

Summary

Our goal is to develop profitable trading systems that use rules that have been derived by learning patterns in price data that precede price changes.

The systems we develop will be quantitative. At every point in their development and use, there will be metrics to help make decisions.

Many trading decisions are subjective. We will use objective functions to quantify subjective preferences.

The system’s characteristics are determined in part by the desires of the trader and in part by what is achievable within the specified risk tolerance. They are also determined by, and to some extent restricted by, the practicalities and realities of combining sound practices of mathematical modeling, simulation, and statistical analysis with uniqueness of financial time series and the business of trading.

George Box famously wrote: “Essentially all models are wrong, but some are useful.”⁷ The more complete quotation adds some qualifications, including: “all models are approximations.” Understanding that trading systems are not perfect, my hope is to help you develop systems that are useful.

---