Dustin Boswell
Core 1b Final Draft

Finding Patterns on Wall Street

Introduction: So you think you can beat the pro's huh?

Most everyone has entertained thoughts of discovering some hidden pattern in the stock market, and then predicting future prices to become rich. And with the digital age upon us, there is little to prevent Joe Shmoe from trying out any wild theory he has thought up. Accurate records of daily stock quotes have been kept for over a century.[1] In addition, PC's are powerful enough to run simulations of what you "would have made" trading by a given strategy. So it seems that the key ingredient for a successful trading strategy is the insight to discover it.

Now obviously, not everyone is making money. In fact, even professional money managers are praised when they do better than the S&P500 (a simple average of stock prices that is used as a benchmark). So then, is *anyone* beating the street – consistently, that is? Surely if you received the "Early Edition"[2] of the newspaper, making 1000% returns yearly would be a trivial task. Realistically, financial wizards making much more than 25% in the long term are the few and the proud. And the Wall Street Journal has often reported on how random selections of stocks compare against those chosen by prudent investors.[3] The horrid implication is that a blind monkey throwing darts at the paper could pick stocks as well as you could. Surprisingly, there is much theory behind

[1] See www.crsp.com for example, a company selling such data.
[2] A CBS drama where the main character mystically receives the news before the events happen.
[3] In http://www.wsj.com/public/current/articles/SB907719524729880000.htm the Wall Street Journal reported an earnings of 10.9% (pros) vs. 4.5% (darts) for 1990-98. However, Liang, in the *Journal of Finance and Strategic Decisions* noticed excess profits did not occur for holding periods greater than a week, and that the excess profits could be attributed to analyst recommendations affecting prices during the short-term holding period (see http://www.tmag.com/jfsd/pdffiles/v8n1/liang.pdf).

why. The Random Walk Hypothesis has long been held as a description that stock prices move randomly from day to day. Closely related, the Efficient Market Hypothesis claims that the markets are "efficient" and that if there were an opportunity for easy profit, then it would have been exploited immediately, eliminating the opportunity in the process. But for those of you who consider your stock-picking abilities superior to that of a low-order primate, the next two sections will reveal the truths and myths of these theories.

Let's take a random walk.

The Random Walk Hypothesis claims that stock prices essentially follow a random walk (possibly with a drift in one direction). That is, although stocks tend to go up on average in the long term, their changes from day to day are random and cannot be predicted from past movements. Naturally, many day-traders watching charts intently for signs, won't hear of it.

Although the claim seems more like a descriptive one, (i.e. the markets *seem* to be random, and we can *test* if they are random, etc…) there is actually much theory for why stock prices *should* move randomly. One of the earlier works on the subject was due to Samuelson (a Nobel Prize winner in Economics), who wrote an article in 1965 titled "Proof that Properly Anticipated Prices Fluctuate Randomly." Note that the Random Walk Hypothesis is based on much of the same reasoning and assumptions as the Efficient Market Hypothesis, namely that all investors have access to the same information, and that they are rational profit-seekers.

And for a long while, it was very much the consensus that stocks did follow a random walk. In sharp contrast to this theory, however, there seemed to be much empirical evidence of statistical anomalies in the past. The well-known "Monday effect"[4] is a good example. This interesting peculiarity of stock prices in the last century shows that stocks have typically gone down on this day, and that trading accordingly would have roughly doubled your returns – a regularity that should occur with very low probability (.002) if the prices were in fact random.

There has also been a "January effect" by which stocks have gone up during this month. It has also been found that stock returns are positive around the first half of calendar months,[5] and unusually high returns were noticed during the Christmas to New Years period.[6] These phenomena occurred for nearly a century. And the list of further "calendar effects" goes on. Is it possible that such simple patterns exist on Wall Street, and that if those mentioned don't still exist, there are others waiting to be found?

Finding patterns in the stock market is a tricky business. The problem isn't that it is *hard* to find good patterns in past stock prices, but that it's *too easy*! Suppose that after leaving your pattern finder program running overnight, you discover a trading method that would have given you huge profits from 1930 to the present. What can you conclude? Those of you familiar with the important difference between in-sample and out-of-sample data may have a smirk on their face. The answer is: *nothing*. From a mathematical standpoint, you have to test your trading method against a period of stock quotes that were *not* used in finding this "perfect strategy". And even so, the exact formulation of how to compute confidence intervals given a certain size data and test set is very rigorous, and cannot be estimated by "gut instinct."

---

[4] See Sullivan, Timmermann, and White, "Dangers of Data-Driven Inference: The Case of Calendar Effects in Stock Returns", June 1998 for an overview of the various calendar effects along with references to the original authors discovering them.
[5] Ariel, R. A., 1987, A Monthly Effect in Stock Returns, *Journal of Financial Economics* 17, 161-174
[6] Lakonishok, J. and S. Smidt, 1988, Are Seasonal Anomalies Real? A Ninety-Year Perspective, *Review of Financial Studies* 1 (4), 403-425.

Surprisingly, given a set of completely random data, it is a fact that one can easily fit a "pattern" to the set. The problem is known as *overfitting* the data. The issue arises when using an overly rich paradigm of patterns to fit the data with. When the size of the set of patterns becomes comparable to the size of the different permutations of data, the problem becomes obvious. For example, imagine a city of people who all think they have a sure-fire way to predict lottery numbers. In a large city, we expect that someone does in fact win the lottery. But if that lucky winner came to you claiming that her method was the reason she won, and that she could predict the next week's numbers, you wouldn't believe her. But why not? The odds of her system winning (assuming lottery numbers are truly random) are 1 in 14 million. While this is true, the odds of *some system* winning on a given week are quite high. This is because there are so many theories out there, that one of them is likely to appear true.

While this point seems obvious when drawn out in the preceding manner, a similar problem fooled many researchers into concluding that there was in fact a Monday Effect. The study concluded that simply being "out of the market" on Mondays resulted in more than above average gains, and that the P-value of such a system (the probability of it occurring given that the data was in fact random) was tiny (.002). When later reviewed by Sullivan, Timmermann, and White[7], it was noted that the probability of *any* successful calendar effect occurring (in the universe of day of the week effects, month of the year effects, etc…) was quite high (around 50% in some cases). That is, we should expect a significant calendar effect to occur about every other time period. So while this does not preclude the existence of patterns in the stock market, it does heed a warning

[7] Sullivan, Timmermann, and White, "Dangers of Data-Driven Inference: The Case of Calendar Effects in Stock Returns", June 1998.

that patterns are a dime a dozen, and care should be taken to realize their true significance.

There have been many statistics done to determine if stock prices do fit a random model. In Lo & Mackinlay's A Non-Random Walk Down Wall Street, their research rejects (at common significance levels) the claim that weekly stock prices move randomly. The randomness of monthly data, however, could not be rejected. It should be noted that the various mathematical tests used to determine whether stocks have moved randomly are quite complicated (perhaps the reason why there was so much debate). Even so, it seems the general consensus that in the long term, markets (for all practical purposes) do follow a random walk, while at high frequency observation it appears to be the opposite.

Wall Street: the Efficient Machine.

The Efficient Market Hypothesis is closely related to the Random Walk Hypothesis, yet it has different implications. It can be stated as "the current price of stocks fully reflects all public information at the time." It strikes a fatal blow to any of us hoping to find a simple way to make big money in Wall Street. Basically, if there was a way to predict future prices from past data, then it would have immediately been exploited and eliminated. For example, suppose it were the case that when a stock with very high earnings goes down three days in a row, it has a high chance of rising the forth day. The theory maintains that many investors would soon catch on to this pattern and be so eager to buy the stock on the third day, that the price would be driven up by demand, thereby destroying the opportunity.

The claim that markets are efficient is a normative one, rather than a descriptive one (although there is little data to refute it). It is a stronger theory than the Random

Walk Hypothesis, and market efficiency need not imply that stocks move randomly.[8] Rather, efficient markets are often modeled as "Martingales", which are a generalization of a random walk. Formally, if x_t is the price of a stock at time t, and I_t is the information at that time, then a martingale process would have it that

$$E[\, x_{t+1} \mid I_t \,] = x_t$$

For example, suppose you know that a certain stock will either go to $105 or $95 tomorrow with ½ probability each. How much would that stock be priced at today? If the markets are efficient, it would have to trade at $100, or else one could buy (if the price were lower), or sell short (if it were higher), and make an expected profit.

The Efficient Market Hypothesis means bad news for anyone trying to make a quick buck. And unfortunately, the data seems to indicate that markets have acted efficiently in the past[9]. Yet after closer examination, the theory starts to contradict itself. "Two economists are walking down the street, and the one sees a hundred dollar bill on the ground. As he goes to pick it up the other one stops him saying 'if it were a real hundred-dollar bill, someone would have already picked it up.'" This sort of logic has been used to make the argument that "perfectly informationally efficient markets are an impossibility."[10] In fact, many behavior models based on game theory, like the Efficient Market Hypothesis, lead to contradiction when taken to the logical extreme. Take, for example, the "beauty contest" game where a group is asked to pick a number from 0 to 100, with monetary incentive for guessing the number that is closest to 2/3 of the average of the groups' guesses. Everyone will soon realize that the average will be in the range

---

[8] Lo & MacKinlay reffering to Leroy, S. F., 1973, "Risk Aversion and the Martingale Property of Stock Returns," *International Economic Review* and to Lucas, R.E., 1978, "Asset Prices in an Exchange Economy," *Econometrica.*

[9] See for example White, H. "Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns."

[0,100], and that 2/3 the average will be in the range [0,66]. And so it would be foolish to guess a number outside of [0,66]. Yet realizing that everyone in the group will soon realize this, the same argument can be made again, and again, until the rational guess would be to choose 0. Yet in practice, we see that the average is around 37 [11]. One could draw an analogy between the average of the guesses, and the inefficiency in the market: although it is not zero, it must remain small if people are prudent enough.

Nevertheless, most experts agree that markets have some, but very small inefficiencies. The reasons arise partly from two of the key assumptions of the Efficient Market Hypothesis. First, it assumes investors are mathematically correct, unbiased profit-maximizers. Second, it assumes that all investors have the same public information at the same time. In practice, if one had a means to acquire information before others did (either illegally with insider information or by some technological advantage), then the theory breaks down, and "easy" profits can be made.

Behavioral Finance: the not-so-rational investor

Much of the formal theory of markets and economics presupposes that investors are rational. Although there are concrete attributes of the rational investor, they can be summed up in saying that the rational investor is one which always seeks to maximize his or her profit; that he or she would never execute a trade unless it is mathematically better to do so. Although we would assume that intelligent, well-educated individuals would act so, empirically it is not always the case. These points are raised not only to cast doubt

---

[10] Lo & MacKinlay in referrence to Grossman, S., 1976, "On the Efficiency of Competitive Stock Markets where Trades have Diverse Information," *Journal of Finance* and Grossman, S., and J. Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets," *American Economic Review*

[11] Nagel, R. (1995), "Unravelling in guessing games: An experimental study", American Economic Review, 85, 1313-1326.

on theories such as Efficient Market Hypothesis, but also to point out the potential success of a method that can correctly identify and react to such irrationalities.

Most of the emotions that compose the supposed irrationalities are of two types: greed and fear. It is almost self-explanatory that when humans are dealing with large amounts of money that can easily grow or disappear, emotions might come into play. As an example, "Some economists have concluded that investors typically consider the loss of $1 dollar twice as painful as the pleasure received from a $1 gain."[12] One experiment showed that people's risk-aversion changes depending on the "framing" of two identical situations. Group 1 was asked "assuming you were just given $1,000, choose between A) a sure gain of another $500 or B) a 50% chance to gain $1,000." Group 2 was asked "assuming you were given $2,000, choose between A) a sure loss of $500 or B) a 50% chance to lose $1,000." In the first group, 84% chose A; in the second group, 69% chose B.[13] Both scenarios, however, are mathematically equivalent. This suggests that people consider more than the necessary mathematics of a scenario.

There is also much study on the "fear of regret". It seems that the regret and remorse experienced when selling a stock for less than it was purchased affects the way investors act. Investors are reluctant to sell losing stocks and "admit" to the mistake, and are eager to collect profits from winners. It has been claimed that investors are more likely to buy "popular" stocks because the regret of otherwise "missing out" would be high, and the comfort of "well, we *all* did bad…" is more present than for lesser known stocks.

Overconfidence is another problem that plagues most investors from correctly assessing uncertainty. When asked to give a 98% confidence interval of the range that

[12] http://www.investorhome.com/psych.htm

the Dow Jones would be in the next month, most investors were not able to do so, and instead gave the equivalent of an 85% confidence interval.[14] Similarly, research has shown that when an analyst is 80% sure a stock will go up, he turns out to be right only 40% of the time.[15] Interestingly, it appears that meteorologists and handicappers at racetracks are the most calibrated in their estimation skills.[16]

Investors also have the tendency to be sentimental when it comes to buying and selling stocks. For example, individuals are more likely to invest in a company that is from their hometown, or that a relative works for, and are more reluctant to sell these stocks.

Thus there are many reasons why markets aren't perfectly efficient. And given this, the question becomes "how inefficient are they?" Or more directly put, "is there enough inefficiency for me to exploit it for a profit?"

So how do people go about finding patterns?

The most typical "pattern finding" techniques are those of watching moving averages along with current stock prices to detect trends as they emerge. Most online brokers and trading software come equipped with tools to analyze such trends. From an Efficient Market standpoint, this is a bad thing since all this "information" is essentially public, and so methods using this metric are less likely to yield big profits.

Along the same line of thought, the more advanced a technique, the lesser known it is to the average trader. One such advanced technique is using *neural networks* to

---

[13] Daniel Kahneman and Amos Tversky, "Prospect Theory: An Analysis of Decision Making Under Risk," Econometrica, 1979.
[14] Daniel Kahneman and Mark Riepe, "Aspects of Investor Psychology," Journal of Portfolio Management, 1998.
[15] Eaton, Douglas R. "The Psychology Behind Common Investor Mistakes," AAII Journal, April 2000.
[16] Daniel Kahneman and Mark Riepe, "Aspects of Investor Psychology," Journal of Portfolio Management, 1998.

discover price patterns. Although they are becoming more and more popular, the good news is that most people don't know how to use them properly.

Neural Networks for Dummies:

Suppose you are convinced that the daily stock statistics (high price, low price, bid/ask spread, volume, etc…) contain predictive information of the next day's price movement. And now you sigh as you sit down to write a program that goes through all the combinations of using volume vs. high price vs. etc… to predict tomorrow's price. Or worse yet, you attempt to describe tomorrow's price as a weighted sum of those statistics (worse because linear methods are unable to discover more complicated patterns). Basically, you just want to know if there is any combination of the statistics that has predictive power. The neural network is your solution.

From a "black box" perspective, a neural network is a construct with many inputs, which take the data you have, and an output, which (hopefully) gives the unknown you are trying to determine. In our case, the inputs would be the daily stock statistics of a certain stock, and the output is the next day's stock price. Inside the box are a large number of interconnections and other *neurons*, all of which have an associated *weight*. The user however, needn't be concerned with these weights. The user merely trains the neural network by presenting past data along with the "correct" answer so that the network reconfigures its internal weights to "learn" from that example. Given enough data, the network will wire itself so as to model an interpolation of the data set. That is, if there was some underlying pattern common to the individual stocks in the *learning set*, the network will attempt to learn it. (It should be noted that since neural networks implement a *non-linear* function of the inputs, they are much better suited to learn chaotic data sets (like stock prices) than any linear "weighted sum" or combinatorial method.)

In practice, it is vastly more complicated. The most plaguing problem is that of *over-fitting* the data, as mentioned before. Neural networks, much like humans, are quite happy to find patterns in any data you give them. And unless you have gone through the mathematical precursors of determining how much data you need, how many neurons should go into the box, etc… you have little clue as to what (if anything) the network has learned. Setting up neural networks is a deep problem, on which many books have been written. This is not meant to serve as discouragement from neural networks, as they are a powerful pattern-finding paradigm. However, the enthusiastic reader should heed the warning that most beginners will become prematurely ecstatic when they find their network has "learned" to predict prices over the past century, only to realize later that it has merely *data snooped* in much the same way calendar effects were found.

Nonetheless, the prospect of designing a network that truly finds a pattern (perhaps because of some unique set of inputs) is real. Regardless of the paradigm used, it is the case practically that some people claim they have a good trading system, yet cannot make a profit from it because the profit margins made on the frequent stock trades are so small that commissions (paid on each trade) make it unprofitable. If indeed transaction costs are the reason that profits cannot be realized, then it is important to find ways to *leverage* the small stock movements predicted by the trading system.

Options 101:

Options are an advanced tool that investors can use to both hedge risk and act on speculation. There are many types of options and combinations thereof, but for our purposes we will focus on the two basic types: *calls* and *puts*. The best way to explain a *call* is through an example.

The Call:

Suppose you are at the supermarket and seedless watermelons are priced at 10 cents a pound (which you feel is a good price). However, you can't (or don't want to) buy them today, so instead you ask for a rain check. This rain check is simply a piece of paper which states "Mrs. Smith is entitled to purchase up to 20 pounds of watermelon at the price of 10 cents per pound up until March 1st." Notice that you aren't *obliged* to purchase them, and if the price of watermelon happens to be even cheaper the next week, you certainly won't *want* to purchase them at the old price. That is, you have the *option* to *exercise* the rain check.

A particular call option on (say) IBM stock would be very similar to this. A realistic example might read "John Q. Investor has the right to purchase 50 shares of IBM stock at the price of $80 per share, up until the last Friday of April." Now in a supermarket, they might offer you a rain check free of charge, as a courtesy, but in the world of finance investors must pay a premium to obtain such a call option. *Why would any investor pay for such a thing?* Well, let's say you are fairly confident that the stock of IBM corporation (which is currently $80 a share) will rise to $90 in the next couple weeks. But you have money tied down in a CD that won't become available until after that time. Let's say you purchase the call option for $2. In two weeks (when the price is $90 and you have the money to buy it), you can exercise your option (buying it for $80) and turn right around and sell it (at the market price of $90) to make a profit of $10. Notice that you turned $2 into $10 without necessarily having the capital of the underlying stock (that is, you could have borrowed the $80 from the bank for a day, which is very cheap to do). It should be noted however, that if the stock had gone down to $75, you would not have exercised it, and you turned $2 into $0.

The ability to leverage your money in this way is one of the big appeals of using options. Suppose you had purchased the stock outright (in the previous example) for $80. If the price increased to $90, you only made 13%, compared to the 500% you made with options. Needless to say, had the stock gone down to $75, purchasing the stock outright would still leave you with $75 worth of stock (a loss of 6% percent compared to 100%).

In the previous example of the IBM call option, the $80 is the *strike price*, and in general, call options exist for a wide range of strike prices (which needn't be equal to the current stock price). You could purchase an IBM call option at a strike price of $120. *Why on earth would you pay for the right to buy something at a higher price than it currently is?* Well, not many people would, which is why the premium for such an option would be very low, maybe 20 cents. But in the rare event that the stock price soars to $130, you would have turned 20 cents into $10 (an enormous return of 5000%).

The Put:

A *put* option is similar to a call, but instead of having a right to *buy*, the option gives you a right to *sell*. Another example will explain its usefulness.

Suppose the current stock price of the company *priceline.com* is $200 and you think the price might go down in the future. Imagine you obtain a put option which reads "Theodore Bear has the right to *sell* 500 shares of priceline.com stock at the price of $180 until the last Friday of January" for which you paid a $5 premium. Then suppose the unimaginable happens, and the stock crashes down to the price of $20 a share. In that case, you would purchase the stock at the current price of $20 and sell for (your option's strike price of) \$180. You have just made \$160 from \$5. Once again (just as the call example), if the stock goes the opposite way that you hoped, the option will be worthless.

The field of options trading is very rich and complicated, and in many ways, the above introduction is an oversimplification. The points to be taken are these:

- Calls can be used to make money when a stock goes up; Puts can be used to make money when it goes down.
- Options allow investors to leverage their capital; in exchange for greater risk, one can possibly achieve greater profits.

In the given examples, the premium price for options was given hypothetically, but the reader may wonder *"Who is selling these options and setting the prices?"* The person who initially creates and sells the option is known as the *writer*, and is typically an investment bank of some sort. And since they too are seeking to make money, they sell the options at a price that is likely to make them money. You would buy an option if the probability of it making money combined with *how much* money it can make is larger than the price of the option. (This is a *gross* simplification.) When you exercise an option, it is the writer that must see through on the other end. That is, the profit you make when exercising a profitable option is the loss that the writer must endure. Thus, the writer will want to sell the options at a price that is likely (on average) to make him money.

Option pricing is not an exact science. For certain types of options (namely the American-style options which are the most common), there is no closed-form solution for what the optimal price (from a writer's point of view) should be. There are certain types (the European-style options where investors can only exercise options *exactly on* the expiration date) where there is an explicit formula for the price. However, this formula is a function of many variables, the most important of which is *volatility*. Intuitively, this is easy to understand. If a certain stock is very volatile, and its price is likely to swing by 50% in a given month, then a call option (whose strike price is near the current price) is likely to make the investor money. And so the premium for such an option would be high. Similarly, the premium for an option for a stock with lower volatility should be lower since it is less likely to be exercised.

Volatility estimation, however, is another difficult task. By predicting the volatility of a stock for the next month, you are implicitly predicting the range of what the price will be (which is one step down from predicting the actual *direction* of movement). Needless to say, if you could predict volatility with any certain precision, you would soon be wealthy. Nevertheless, option-pricing is not exact, and if one could devise a method to price them more accurately than the current methods being used today, easy profits could be made by purchasing options that are under-priced.

But an interesting property of martingale processes (by which stocks move according to the Efficient Market Hypothesis), is that volatility is "allowed" to be predictable. Going back to the original example, if XYZ stock was known to be either $95 or $105 tomorrow, each with probability ½, then the price today should be $100. But if you knew it would either be $75 or $125, the price should be exactly the same, since the *expectation* of XYZ's price has not changed. The fact that the Efficient Market Hypothesis does not preclude the forecastability of such "higher order" data has gathered interest in volatility prediction methods.

Thus options offer the speculative investor incentive in many ways. If an investor has a trading strategy that theoretically makes money by buying and selling stocks in the short term, but practically is obstructed by transaction costs that eat up the small percentages gained from trading, options may provide the leverage needed to exploit small gains into reasonable profits. More generally, if an investor has a system that can predict any sort of data (price, direction, volatility, etc…), then that in turn can be used to price options more accurately and take advantage of option price discrepancies.

Conclusion: There's no such thing as a free lunch

Some people are enthralled by the possibility that underneath all the apparent chaos that is Wall Street, there is some divine order that explains it, just waiting to be discovered. Although the field of financial prediction furies with debate, there is much agreement that the existence of a model which can fully predict the future of the Dow Jones is not possible.

Stocks appear to move randomly in the long term, while in the short term it is not the case. And the little (if any) inefficiency is most likely to occur for small trading periods, during which the assumptions of information symmetry made by the Efficient Market Hypothesis may break down. Combined with the other slightly inaccurate assumption of investor rationality, this may provide the attentive market follower with opportunity for profit. It is quite possible therefor to make a living trading stocks. The general consensus however, is that it requires acting on information that has not yet reached the public (before others do, or better than others do). So the old cliché remains true: there's no such thing as a free lunch; you still have to work for it.