Chapter 8: The Science of Strategy Evaluation

The science of strategy evaluation has two basic parts to it. The first part is the evaluation of the financial aspects of the strategy. How do we measure profitability? Is a particular trading strategy a better place to put your money than alternative investments or businesses?

The second part of strategy evaluation is more personal in nature. The strategy must be evaluated in light of the person who will be doing the actual trading. This is what I call statistical evaluation. Does the historical performance make this strategy acceptable to the personality and trading style of the individual trading it? Does this trading strategy have characteristics that will allow the person to trade it effectively and have the discipline to execute it? Will the trading of this strategy provide too much emotional stress? The statistics will tell us.

And finally, it is important to know when your strategy has stopped working.

Financial Evaluation

There are two ways to evaluate a strategy financially. First a strategy may be evaluated on its own merits as compared to alternative forms of investing. That is, the return on invested capital over a period of time. How does the particular trading strategy stack up as compared to T-Bills, common stocks, etc?

Second, a strategy should be evaluated financially on its own merits. This means is it viable as a trading strategy as compared to other trading strategies? Does Strategy A provide a better return than Strategy B?

RISK-FREE RATE OF RETURN

The place to start when contemplating a trading strategy is with the risk-free rate of return. This is the return you would expect to receive on an asset that is virtually risk free. Most analysts use the 90-Day US Treasury Bill rate as the risk-free rate. And while it could be debated as to whether the debt of the US Government is risk free, it is a close as we can come.

The first and most obvious principle is that any strategy must provide a greater return than the 90-Day T-Bill rate, or you would simply be better off just putting your money in T-Bills.

However, you must also assess the return that you will require of the strategy in order to compensate you for the added risk. How much income over and above the T-Bill rate is required to entice you to take your money out of T-Bills and put it into a trading strategy? You should assess the premium that you will require for trading a particular strategy.

As the risk is greater for trading stocks and futures, this premium should be quite large. I have always recommended that for stocks you should at least double the T-Bill return rate, and for futures you should require four times the T-Bill return. If the T-Bill rate is 6%, I would require at least a 12% return per year for stocks and at least 24% per year for futures before I would consider taking my money out of T-Bills and putting it in these markets.

If the current T-Bill rate were 10%, I wouldn't be interested in a strategy for futures that did not return at least 40% per year. If the historical testing did not indicate that this 40% return was possible, I would keep my money in T-Bills.

Ultimately, you must determine your own risk premium. Take some time to think about what you consider to be a reasonable return for your trading efforts. It might not be my four times the T-Bill rate; you might only require three or two times. But if you are not compensated for the increased risk, it is more prudent to place your money elsewhere.

Also note that using my recommended approach permits the required rate of return to change over time in that there have historically been large swings in the T-Bill rate. In times of high inflation, like during the late '70s, the T-Bill rate generally rises, thus requiring a higher return for your trading account. In times of

low inflation, the T-Bill rate lowers and therefore you would not require as high a return from your commodity or stock trading accounts.

In times of high inflation, the volatility of most stocks, commodities and futures increases, thus providing the opportunity to profit from this increased volatility. You must make sure that that your strategy will provide the necessary return in different financial environments (high or low inflation, recession, etc.).

THE INVESTMENT

I believe that the decision to place money in a trading strategy either for futures or stocks should be made with the same due diligence and financial analysis as putting your money in any investment. What is the expected return on investment? What are the relative risks to achieve the expected return?

Creating a strategy and running historical tests is no different than investing in real estate, leases, mortgages or even junk bonds. All investments are sold to the public by presenting the expected rate of return and estimated risks. These returns are evaluated by looking at the past history of the investment and making assumptions that the future will be similar to the past.

Devising a trading strategy and implementing it is no different. We run historical tests and make assumptions that the future will be similar to the past. Although we intuitively know that it won't be exactly the same, we make the assumption that it will be close enough to induce us to risk our funds on this strategy.

Before we can calculate a return on investment figure, we need to determine just what our investment is. The place to start for futures is with Maximum Intra-Day Drawdown (MAXID).

I consider the MAXID to be my investment in my strategy. If you were to operate any type of business, you would have to invest money in facilities, inventory, and labor before any revenue came in. Then, you would calculate your profits as a percentage of this investment.

THE RETURN ON INVESTMENT

In trading strategies, to evaluate the return on investment I use what I call ROMID—Return On Maximum Intra-day Drawdown. I view MAXID as the investment and calculate my return based on this number.

Futures margin should not be included in the calculation of the investment for three reasons. First, since it is now standard practice to keep margin in T-Bills, it

is redundant to use margin in the calculation for ROMID. We should either eliminate the margin or include the interest earned on the T-Bills to calculate the return on investment. I eliminate the margin.

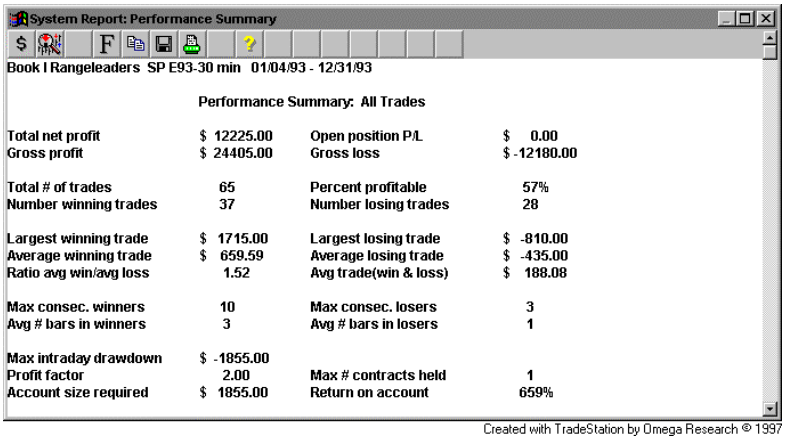
The second reason not to include margin as part of the investment is that margin requirements change frequently. This would force us to estimate an average margin over a period of years, which would distort the year to year returns. In the case of the S&P futures, for instance, this would be difficult given the wide swings in margin over the last 10 years.

Third, using ROMID facilitates the comparison of different strategies on different futures contracts. Comparing the Return on Maximum Intra-day Drawdown eliminates the differences in margin, concentrating on the return for actual funds at risk.

It is for these reasons that I do not recommend that you include margin in your calculation of the investment. Use MAXID as the investment and ROMID as the return on investment. This will facilitate the very important process of comparing returns on many different strategies.

Statistical Evaluation

PS 1 is a Performance Summary, which we have been looking at throughout this book. This is the financial information for the strategy and is the trader's equivalent of a corporation's Balance Sheet. This is the basic information that we use to analyze and compare trading strategies.



System Report: Performance Summary

Book I Rangeleaders SP E93-30 min 01/04/93 - 12/31/93

Performance Summary: All Trades

Total net profit	$ 12225.00	Open position P/L	$ 0.00
Gross profit	$ 24405.00	Gross loss	$ -12180.00
Total # of trades	65	Percent profitable	57%
Number winning trades	37	Number losing trades	28
Largest winning trade	$ 1715.00	Largest losing trade	$ -810.00
Average winning trade	$ 659.59	Average losing trade	$ -435.00
Ratio avg win/avg loss	1.52	Avg trade(win & loss)	$ 188.08
Max consec. winners	10	Max consec. losers	3
Avg # bars in winners	3	Avg # bars in losers	1
Max intraday drawdown	$ -1855.00		
Profit factor	2.00	Max # contracts held	1
Account size required	$ 1855.00	Return on account	659%

Created with TradeStation by Omega Research © 1997

PS 1

Sample Performance Summary

There are basically two parts to statistical analysis. The first are the statistics that reflect the viability of the strategy itself. The second are the numbers that are crucial when considering whether or not you can actually trade this strategy.

STRATEGY VIABILITY

The four numbers that are statistically important and reflect the viability of the strategy are: Total Number of Trades, the Average Profit per Trade, the Largest Winning Trade, and the Profit Factor. If these do not pass our initial test, then we look no further and try another strategy. However, if these four pass our minimum requirements, we then look at other values in the Performance Summary to see whether or not we could actually trade this strategy.

Total Number of Trades

The first number we look at is the number of trades. This should be a statistically significant number. The basic rule is the more trades the better. Ever since I started trading, 30 trades has been bantered about as the number of trades per signal required for a strategy to be statistically sound. I am not a statistician so I can't comment on the validity of this number. Nevertheless, I have always used it as sort of benchmark.

You have to draw the line somewhere and it might as well be 30 trades. The most important thing to remember the less trades in a test, the more skeptical you should be about the strategy's performance in the future. If I produced a strategy that had 200 trades and compared it to one that had 25 trades, I would certainly be more confident about the 200 trade strategy. If I produce a strategy that has 30 trades or less, the red flag goes up and I look at the strategy very carefully.

Average Profit per Trade

As another initial filter, I use the average trade (average profit per trade). It is this number that tells you how much room you have for trading mistakes. Even if you use a high number for slippage and commissions, you must have enough latitude in the average trade to cover several more ticks of slippage. You simply do not want to underestimate the possibility of greater slippage. I always want at least $200 per trade as an average, after slippage and commission. This ensures that even with a few more ticks of slippage there will be enough room for profits.

Largest Winning Trade

The Largest Winning Trade is a significant number as it relates to the Total Gross Profit and Net Profit. The issue is that if a large portion of the profits of a strategy come from one trade, we have a major problem with the strategy. I recommend that the Largest Winning Trade be no more than 50% of the gross profit or 25% of the Net Profit.

For example, over the years I have seen many profitable trend-following strategies on the S&P futures. Quite often, however, upon closer scrutiny I have found that most, if not all, of the profits have come from one short trade during the 1987 crash. If you take this one trade out, you would see that it distorts the profitability of the strategy and the profits would be dismal.

Profit Factor

The Profit Factor is calculated by dividing the Gross Profits by the Gross Losses. I view this amount as the risk/reward ratio. That is, how much reward am I going to get for risking $1.00?

My personal level is 2 to 1. I always want to at least have a 2:1 risk/reward ratio. If the Profit Factor is not greater than two, I will usually not trade the strategy. I work very hard to get a Profit Factor greater than 2.

I also use the Profit Factor to compare strategies. Most traders will look at the Net Profits or the ROMID to compare the effectiveness of a strategy. It is logical; the most profits or the most return on maximum intra-day drawdown. But for me, I like to look at the strategies with the greatest risk/reward ratio. The Profit Factor always clears up any ambiguity I might have when the Net Profits and the ROMID of several strategies are very close. In this sense, I use it as a tiebreaker.

PERSONAL EVALUATION

The second part of the statistical evaluation has to so with the characteristics of a strategy that have a bearing on your ability to trade it. If a strategy passes the financial test, and makes it through the first four statistical filters, then we are ready to look at the trading statistics to see if the strategy fits our personality and risk profile.

MAXID

The first number I usually look at for a trading strategy is the Maximum Intra-Day drawdown or MAXID. This number tells me the actual and maximum cash outlay that will occur at any one time to support the strategy. This is the major cost of doing business, in other words, the cost of maintaining the strategy.

MAXID calculates the amount of money that it takes to sustain a drawdown of funds between two new equity highs—losing trades plus slippage and commissions. It is the maximum amount of funds that you need to give up to invest in the market to get to the next new high in your account. This is what I call your pain threshold.

Why is this amount important?

First, you need to be financially able to withstand this kind of dip in your account. If this dollar amount is a stretch, then you should either find another strategy that has a lower MAXID or put the money in T-Bills.

Second, even if you can financially withstand the drawdown, the real issue is whether or not you could psychologically stand the pain. I know many traders who design strategies with a small drawdown because they are simply unable to take sustained losses. That's perfectly acceptable. Remember the whole psychological key to trading is to be able to take the losses. If you are uncomfortable with the level of MAXID, then you should find or design a strategy that has a level that is comfortable for you.

Percent Profitable Trades

As you look a the percent profitable, you have to ask yourself whether you can live with a strategy that has less than a 50% win rate or if your personal trading style requires more positive feedback.

Some traders can psychologically handle 40% or 35% winners. They have confidence in the historical data and know that even with this low rate of wins they will make money over time. Others will not be able to live with this. Being subject to such a large percentage of losers would produce much anxiety, decrease their confidence level, and most likely cause them to abandon the rules that make the strategy work. This is a prescription for strategy trading failure.

On the other hand, having a high percentage of winning trades does not necessarily make a better strategy. Many of the best performing and most profitable strategies I have seen have a Percentage Profitable Trades number in the 35% to 45% range.

So, the Percent Profitable Trades number has no real practical value other than psychological. You should think about this issue and the percent of profitable trades you could live with and would be able to trade effectively.

Maximum Consecutive Losers

This number's importance is again psychological. Just how many losers in a row do you think you could sit through before thoughts of abandoning the strategy enter your mind? 7? 8? 10? 3? Only you can assess this and decide. It is a personal matter, and it is of no practical value other than psychological.

And even if you think you could sit through 7 losers in a row, wait until you are faced with actually doing it. Even with great confidence in a strategy, and the historical data to back it up, this is a very difficult thing to do. When it happens to me, I have to keep reminding myself of two important philosophical points. First, the market will eventually have to facilitate trade and move. And second, that when it does move, my strategy is sound enough to catch the big move. These two precepts are what give me the confidence to go through a string of losing trades without losing my confidence.

So here you want to have a number that you honestly feel you could handle. You also should realize by now that it is possible to have a very long string of losing trades, even longer than the historical test, and still have a well designed strategy. Just be prepared when you actually start trading the strategy.

How to Know your Strategy has Busted

So we've been trading our great strategy real time for a while and it's been working very well, but lately we've been experiencing substantial drawdown and a significant number of losing trades. At this point, we need to make a reality check to ensure that our strategy is still working. We want to make an assessment as to the viability of the strategy that tested out great historically but is now losing money.

The first thing to assess is whether the strategy is catching the moves for which it was designed. If your strategy ever misses a move of the type for which you designed it, the strategy has busted.

A trend-following strategy is designed to lose money in sideways markets make it all back and more in the trend. If your trend-following strategy misses the big move, it clearly has busted. A number of losing trades in a row does not mean the strategy is not working. Missing the big move does.

Volatility breakout strategies bust when the volatility of the market changes substantially and the strategy misses the moves for which it was designed. For example, if your volatility strategy historically had 62% profitable trades and only two losing trades in a row, and you recently had a string of 6 losing trades out of the last 8 (25% profitable), with the last 4 being losses, clearly something has changed. The strategy is missing the moves for which it was designed and you should review the strategy.

Volatility strategies are designed for short-term quick profit trades. They have a high percentage of profitable trades. If there is no follow through on the volatility breakouts and the strategy is not performing up to its historical standards, you should reassess this strategy. If it is not capturing those short volatility pops, then something is wrong and the parameters need to be reviewed.

Excessive drawdown can also be a tip off that something is wrong. If for the last 20 years the MAXID has been no greater than $7,500 and we now have $9,000 of drawdown, the red flag should go up. You need to make sure that something important has not changed. If the drawdown exceed two times that which existed in your historical tests, I would stop trading the strategy until you figured out what is going on. Maybe the market is just in a never before seen sideways phase and you should stick it out. But maybe something else has changed. The key is that an overly large MAXID by historical standard is a good indication that the strategy should be reviewed.

Summary

It is an art to design an effective strategy. Strategy design is a creative process that capitalizes on the ability to synthesize new ideas and creatively put them together into a viable strategy.

In performing my financial evaluation, I want to make sure that the strategy itself compensates me for the increased risk over the 90-day T-Bill rate. If it does not, I would rather keep my money in T-Bills.

For personal and statistical evaluation, I use four key numbers:

1. Percent Profitable. What is my pain threshold?
2. Maximum Consecutive Losers. How many can I stand?
3. Maximum Intra-day Drawdown. What can I afford?
4. Profit Factor. I need a 2:1 risk/reward ratio.

I also make sure that the Total Number of Trades, Average Profit Per Trade, and Largest Winning Trade are within acceptable parameters.

The other analysis of the strategy is when we are actually trading it. Is it performing up to its historical potential? If it is not, we need to have a procedure to decide whether or not the strategy has busted. Obviously we should not be trading a strategy that is not catching the moves for which it was designed.

Strategy evaluation is not an art, it is a science. There is a clear procedure with a definite range of acceptable results. Once you have delineated your acceptable limits for the evaluation, the analysis should become routine. When your results move out of this acceptable range, the strategy becomes suspect. Early detection of a failed strategy is as important to long-term profitability as the design of the strategy itself.