

# HOW TO TEST AND INTERPRET TRADING SYSTEM PERFORMANCE

By Wayne A. Thorp

Many forces are at work when you trade a system—commissions, slippage, protective stops, idle interest, margin, and short trading can all significantly influence a trading system's performance results.

Pick up any technical analysis trade magazine, and inevitably you will run across companies and practitioners marketing technical analysis trading systems. Like any other type of investment strategy or methodology, a popular way to determine how one system stacks up against another is by comparing annual returns. While these numbers are helpful in separating the winners from the losers, it is important to keep in mind that a multitude of factors impacts the performance of any trading system.

When judging the efficacy of a system's reported performance or the performance of a system you create, keep in mind several issues:

- Are the performance figures based on backtesting or actual trading?
- Is the system optimized and, if so, how does it perform over "hold-out" periods?
- How does it handle income reinvestment?
- Are there any tax implications?
- What are the assumptions inherent to the system itself—commissions, slippage, and money and risk management stops?

This article will walk you through a general discussion of how these elements can impact the financial performance of a trading system.

## ACTUAL TRADING RESULTS?

When confronted with the results of a trading system, your first thought should be: How were these results generated? If a system claims returns of 25% a year, is this based on actual trading or historical backtesting?

Backtesting involves testing a system using a set of historical data. Results based on actual trading have a greater degree of credibility because returns are generated over actual trading conditions *as they happen*. Secondly, results based on backtesting are more easily manipulated to generate the highest possible return (the practice is called optimizing).

## GAUGING PERFORMANCE

However, backtesting using historical data is the most efficient manner to derive system performance statistics. Backtesting is the fastest and most popular way to gauge the *potential* profitability of a trading system. The process of backtesting involves running a system over historical data. The end result is system performance statistics that show how the system *would have* performed had it actually been used over that time period. In order to backtest a system, all you need is the historical database.

Ideally, whenever you backtest a system, you want to use a "significant" amount of data in order to capture as many different market phases as possible. The amount of data you will require depends, in part, on the system you are testing—real-time, tick-by-tick systems require several days or weeks of tick data while end-of-day systems will need at least several years of daily data. The bottom line, however, is that the more data you have, the more complete the picture you can draw from your backtesting results.

---

*Wayne A. Thorp is assistant financial analyst of AAll.*

A drawback to historical backtesting is that results are based upon events that have taken place in the past. Therefore, the most you can hope to learn from backtesting is how a system *may* perform. There is no guarantee that what has happened in the past will repeat itself going forward. The usefulness of backtesting lies in its ability to provide insight into how a system may react in various market conditions. Backtesting can often show you if a system works better during trending markets compared to trading (sideways) markets, or vice versa.

You should also keep in mind the period over which a system is backtested. If backtested results cover “odd” periods, this should serve as a red flag for possible manipulation. Companies sometimes only report results for the periods in which the system performed best. If the results are for the period 1992 through 1999, you should ask yourself how the system did during the market downturns of 1991 and 2000. Often, the performance of the system outside the reporting period will have an adverse affect on the overall performance. Ideally, you would like to have system results that cover several market cycles—both good and bad.

A final thought to consider is how a system performed in comparison to a “buy and hold” strategy. The whole idea behind trading a given strategy is to garner greater returns than if you simply bought the stock and held it over the period. If you cannot outperform such a strategy, you need to go back to the drawing board and try again.

## SYSTEM OPTIMIZATION

Optimizing is the process of “fitting” a trading system to a specific set of data. For example, suppose you are using a simple moving average system that generates buy signals when the closing price moves above the moving average and sell signals when the

closing price moves below the moving average line. Optimizing would run the system over the data, testing varying moving average lengths to find the period that netted the largest gain or the smallest loss.

The problem with optimizing is that you are finding the best set of parameters for a fixed period in the past. However, there is no guarantee that the past will repeat itself. While optimizing isn’t necessarily a bad thing, it is easy to fall into the trap of over-optimizing. In the end, you may have a system that performs spectacularly in the optimization period, but falls apart when tested over any other period.

One way to validate or disprove the effectiveness of optimizing is through the use of a “hold-out” period—a set of data over which the system is not optimized. Returning to our earlier example, let us assume you have 20 years of historical data for backtesting. A hold-out technique to follow would be to optimize the system over one half of the data (10 years) to arrive at the optimal moving average period length. From there, you would then test the optimized system over the second half of data. If the results from the two 10-year periods are comparable, you can be more confident that the system will perform in a similar manner over other periods and, most importantly, going forward. If, on the other hand, the results over the last 10 years differ dramatically from the first 10 years, you should begin to question the viability of the system.

## OTHER FACTORS

You should be aware of a few factors that, while today’s software does not take them into account, can affect the overall performance of a trading system.

The receipt or reinvestment of dividends is an issue that is not handled by most technical analysis programs. However, it can have a significant bearing on a system’s performance. If you trade stocks that

pay dividends, the dividend income received will have a positive impact on performance.

Another issue that few, if any, trading system packages explicitly account for is taxes. Depending on your holding period—short-term or long-term—the marginal tax rate on your gains will differ. Those holding an investment for over one year are subject to the long-term capital gains rate of 20%. If you hold an investment for less than a year, gains are viewed as income, which is taxed at your marginal income tax rate. Depending on your income tax bracket, therefore, you would need to generate a higher rate of return to overcome the tax effects as compared to someone holding their investment(s) for more than one year.

## SYSTEM ASSUMPTIONS

When you construct a trading system, the assumptions you make (or fail to make) play a role in how well your system may perform.

These assumptions involve initial equity position, trading on margin, the handling of short trades, commissions, time and price slippage, risk and money management stops, and interest earned on idle balances.

### Initial Equity

The initial equity amount is the amount of money you have in your account before you begin trading. By beginning with a sizable amount of equity, you gain greater flexibility in the form of entering a larger position, which, in turn, can generate larger total dollar gains (or losses).

Typically, by entering with more money, you can stay in the game longer. This is especially true if you plan to short stocks. Short sellers hope to profit from stock price declines by borrowing stock and selling it first, then buying the stock later at a lower price and returning the borrowed shares. When a stock is sold short, your potential loss extends well beyond your initial investment. Depending on who you

ask, you will probably receive different answers regarding the “ideal” equity balance. Ultimately, it is up to you, just be sure you can afford to lose it!

### Short, Long, or Both?

One critical issue involves how to deal with sell orders. When a sell is triggered, you could sell your long position and go to cash, or you can elect to be more aggressive and “double down.” This involves selling your long position and establishing a short position in which you profit if the security decreases in value, but you lose money if the security goes up in value.

### Margin

Margin investing is a delicate topic that investors should understand before attempting. Margin is money you borrow from a broker, similar to a loan, that you then use to buy stocks. You cannot buy all stocks on margin: Those priced below \$5, certain other Nasdaq stocks, and IPOs within a certain period of their introduction are excluded.

Brokers are regulated by the Federal Reserve as to how much credit they can extend to their clients. Currently, you can initially borrow up to 50% of the value of your marginable securities for stocks. For example, assume you have \$10,000 in a margin-approved brokerage account. This means you can purchase up to \$20,000 of marginable securities, with 50% coming from you and 50% from the brokerage. Another way to word it is that you have \$20,000 of “buying power.”

The amount you are able to borrow on margin fluctuates on a daily basis as the prices of the marginable securities rise and fall. If the prices increase, so too does the amount you can borrow. The opposite holds true as well: As prices fall, the value of the marginable securities—your collateral—falls as well. If the value of your margined securities falls below a predetermined minimum level, you will receive a “margin call” from your broker. At

this juncture, you are required to either liquidate part of your existing position or send in more money to bring the value of your account back above the predetermined level; or your broker can sell your securities without calling.

Investing on margin carries with it risks and rewards—it magnifies the effects of gains and losses. Returning to our \$10,000 margin account example, let us assume you buy 1,000 shares of stock priced at \$20. You pay for this transaction by borrowing \$10,000 from your broker and using your \$10,000 from your account. If, in a year, the price rises to \$40 a share, the value of your investment has risen from \$20,000 to \$40,000. If you sell the shares and pay back the \$10,000 you borrowed from your broker (including margin interest—interest charged by the broker for the privilege of using their money), you would have roughly \$30,000 remaining—\$20,000 of which is profit to you.

On the other hand, if you simply use your \$10,000 to buy 500 shares of the \$20 stock, your profit would be roughly \$10,000. In the first example, you would have made \$20,000 on a \$10,000 investment, while in the second you would have made \$10,000 on that same \$10,000 investment.

Just as margin can improve your profit, it can also worsen your losses. If the \$20 stock you initially bought on margin falls to \$15 a share, the investment value falls from \$20,000 to \$15,000. After paying back the \$10,000 you borrowed from the broker, you are left with \$5,000 of your original \$10,000. Without margin, the 500 shares you bought at \$20 would now be worth a total of \$7,500. With margin, you lose \$2,500 more than you would have using only your own money. Be aware, too, that in our examples we did not account for commissions, margin interest, or capital gains taxes, which, as we have discussed, will impact the bottom line.

### Commissions

People tend to forget what a dramatic impact commissions—the fees paid for buying and selling securities through a broker—can have on the overall success of a trading system.

To get a more accurate picture of a system’s profitability, it is important to figure in the commission costs. This is especially important for a system that generates numerous buy and sell signals, which will dramatically lower the profits or increase the losses of a system. Commissions can vary greatly depending on the type of security you are trading and whether you are using a deep-discount broker or a full-service one.

### Slippage

Another element that many traders lose sight of is the fact that you will rarely be able to enter or exit a trade at exactly the same price at which the trading signal was generated. If your system is based on end-of-day data, a buy or sell signal will be generated after the market close. Realistically, your first opportunity to act on the signal is at the open the next day. The difference between the price at which the signal was generated and the price at which your order is actually filled is called slippage. When testing a trading system, it is important to account for slippage; otherwise the trading results are overstated. Some software programs allow you to specify slippage in dollar or percentage terms, while others allow you to build in a time delay between the signal and order execution.

### Stops

Perhaps the most useful tool in developing a trading system is a stop. Compared to commissions and slippage, which are costs associated with a system, stops are more of a system “tweaking” mechanism. Stops are user-defined points where a position is closed out. When a stop is triggered, the position is closed regardless of the current

status of your trading rules. Stops allow you to limit your losses should a trade go against you. The stops you specify in a trading system are similar to stop-loss orders you can place when executing a trade. As the name suggests, a stop-loss order is designed to stop a loss. If you purchase a stock for \$30, you can protect yourself against the possibility of it falling in price by placing a stop-loss sell at \$30. A market order to sell the stock is placed if the stock falls below \$30.

There are several strategies using stops when creating a trading system, the most popular being breakeven, inactivity, maximum loss, profit target, and trailing stops.

**Breakeven stops** close open positions when the closed-out value of the position equals the amount at which the current trade was opened. The stop is placed at the price where the trade could be closed and the proceeds generated would equal the equity value when the trade was opened.

**Inactivity stops** will close an open position when the security's price does not generate a minimum percent or price change within a specified time period. If you specify 1% as the minimum change and 20 as the number of periods, the system would automatically close any long (short) positions where the security's price has not increased (decreased) by at least 1% within any 20-period time frame.

**Maximum loss (max loss) stops** are useful as a risk management strategy, because you can specify the exact percentage or dollar amount of your total equity you wish to risk on a given position. These stops close an open position when the losses resulting from the trade exceed the specified maximum loss amount.

**Profit target stops** exit a trade once it reaches a predetermined profit level. Therefore, if you specify 10% as the profit target, open positions will be closed when they generate a 10% profit (after commissions).

Lastly, **trailing stops** close open

positions when a specified amount of the current open position's profits is lost. Each time a position's profits reach a new high, the trailing stop is moved to a level that allows a specified portion of the position's profits to be lost.

You are also able to specify the number of periods to ignore in trailing stops. For example, if you instruct the system to ignore three periods, the trailing stop will lag by three periods. Therefore, the last three periods' profits or losses will be ignored when determining the current stop level. Such lags are useful in filtering out price swings. However, you need to exercise caution when using trailing stops. They are not designed to limit losses, but to lock in profits.

### Idle Interest

Depending on the type of system you are using, there may be times when you are not in a trade. This means that all long trades have been closed and short trades covered. Ideally, you will be earning some interest on this "idle balance." The interest you might earn is influenced by several factors, including the brokerage firm you use to execute your trades, the cash accounts available, and the size of your account.

### HOW IT WORKS: AN EXAMPLE

Now that you know what to consider when testing a trading system and examining the results in general terms, let's take a look at an example of how these factors can impact the performance of an actual system using historical data. For this article, we used MetaStock 7.0 by Equis International.

Before you can begin testing a system, you obviously need to have a system to test. A trading system can be as simple or as complex as you can imagine—from a moving average crossover system to one consisting of several highly evolved indicators. For our example here, we use a 50-day exponential moving

average (EMA). The exponential, or exponentially weighted, moving average is calculated by taking a percentage of today's closing price and applying it to yesterday's moving average, with greater emphasis placed on the newest price. (To learn about exponential moving averages, refer to the August 1999 *AAIL Journal* article, "An Intro to Moving Averages: Popular Technical Indicators" on our web site.)

With our system, buy signals are generated (and short positions covered) when the closing price moves above the 50-day exponential moving average. Likewise, long positions are closed and short positions are entered when the closing price falls below the 50-day exponential moving average. This system may seem overly simplistic, but it illustrates the elements we have been discussing when evaluating, testing, and optimizing a trading system.

To show how the factors such as commission, slippage, and stops can impact the overall performance of a trading system, we must have a benchmark against which to compare their impacts. Therefore, we begin by presenting a system that, in effect, ignores many of these issues.

Using Walt Disney, we ran our initial test over the 20-year period from November 3, 1980, to October 31, 2000. The only assumptions we made for this test are that we handle both long and short trades and that we begin with a non-margin account balance of \$10,000. We do not account for commissions, slippage, stops, or interest on idle balances.

Running this "sterile" system resulted in a net profit of \$20,603.32 over the period. While the system made money, it fell well short of the return netted by a buy-and-hold strategy. If you had bought \$10,000 of Disney stock at the beginning of the period and sold it at the end, you would have earned \$384,480.56! At this point, it is evident that this system needs some improving before it is ready to be traded in the real world.

Next, we apply our assumptions to the system, individually first and then in combination. We begin by testing our system assuming that we borrowed 20% of our equity on margin. Although federal regulations allow you to borrow up to 50%, we recommend this only for experienced traders who are well-versed in the implications of trading on margin. Trading on margin had a slightly negative effect on this system—we netted \$20,461.44, or \$141.88 less than what we would have earned had we not traded on margin. However, if we had followed a buy and hold strategy using margin, we would have earned an extra \$97,000.

Then we tested the system assuming that we pay a \$15 commission for each trade generated by the system—\$15 for each buy and \$15 for each sell. The 807 buy and sell trades the system generated over the 20-year period cost us \$12,105 in commissions. However the true cost was \$14,101.46 since the money spent on commissions can not be spent on trades which may cost us on profitable trades or save us on losing trades. Obviously, depending on the price you pay for transactions and the number of trades you place, the amount you pay in commissions can vary significantly.

Accounting for slippage, we instructed the system to execute trades at the opening price the day after the signal was generated. This adds a greater degree of realism to the system since signals are not generated until after the close of trading for the day. This “delay” in execution had a tremendous impact on the overall performance of the system—a net loss of \$1,604.27, or \$22,207.59 less than the “sterile” system.

In a system such as this, which is fully invested, idle interest is not much of a consideration. In fact, the only interest we earned on our idle balance was during the first 50 days of the system. Since there was no 50-day exponential moving average during this period, we were not in any trades and we earned \$60.

Lastly, we entered in our protective stops for the system. The two we used were a trailing stop and max-loss stop. Our maximum-loss stop closes a trade if it loses 2% of our remaining equity. Therefore, in essence, we are risking 2% of our equity per trade. Remember, however, that because of slippage, we run the risk of losing more than 2% on a given trade. Our trailing stop risks 20% of our profit while ignoring one period to filter out random price swings. Implementing our stops into the system has a significant positive impact—it netted \$102,050.32, \$81,447 more than the sterile system.

Having discussed all of our factors in isolation and showed how they impact the performance of our system, it is time to see how they work in tandem with one another. Our last test combines all of the assumptions we have covered, and the end result stands in stark contrast to the result we first arrived at. In this case, our system exhausted all of the equity in our account, leaving us with a loss—an ending amount of \$9,999.46. Overall, the system generated 502 trades, which cost us \$7,530 in commissions. Furthermore, our idle balance earned \$268.96 over the 3,630 days the system was out of all trades, due in large part to a lack of liquidity to execute trades. Obviously, this system needs some work before it is ready for actual trading!

#### USER ACTION REQUIRED

What sometimes gets lost in the

discussion of trading systems is the fact that, although they are mechanical in their generation of buy and sell *signals*, most programs are not capable of executing their orders for you. Therefore, the performance of your system is ultimately contingent on whether you execute each and every trade when you are supposed to. The most difficult thing for many traders is not creating, testing, or optimizing a system, it is actually following it in real-time.

Depending on the type of system you are trading, you may have to devote a significant amount of time to monitoring it and executing trades. Intraday systems, those based on real-time or intraday delayed data, may require your undivided attention through the course of a trading day. End-of-day systems, while not demanding the same attention, require daily examination. Therefore, time is another intangible cost associated with following a systematic trading strategy.

#### CONCLUSION

It is clear from our discussion here that many forces are at work when you trade a system. Commissions, slippage, protective stops, idle interest, margin, and short trading all in their unique way influence a trading system's results.

Comparing the results of our initial test where we ignored many of these factors to the results generated when we integrated them shows how important it is take them *all* into consideration when evaluating or testing a trading system. ♦

#### Check the following areas of our Web site for more information:

- Use the Search tool to locate related *AAll Journal* articles:
  - “An Intro to Moving Averages: Popular Technical Indicators”
  - “Using Moving Averages in a Systematic Trading Strategy”
- *Computerized Investing* subscribers can access these articles:
  - July/August 2000 Feature: “Building & Testing a Trading System”
  - May/June 2000 Comparison: “Technical Analysis Software”
  - Nov/Dec 2000 Comparison: “Web-Based Technical Analysis Services”
- Go to AAll's list of Top Web Sites under Investor Resources. Scroll down to the best sites for charting/technical analysis to connect to sites that offer trading systems and strategies.
- Download trading shareware from the Stocks/Technical Analysis area of the Download Library.

**AAll.com**  
American Association of Individual Investors