A General Methodology for Testing the Performance of Technical Analysis on Financial Markets

Jérôme Senanedsch[1]

Abstract

Our study provides a general methodology for testing the performance of technical analysis on any market. This methodology is very useful to evaluate whether or not a market is efficient (EMH) in its weak form. The main difficulties of this area of research are encountered when one wants to provide a general result on different tools of technical analysis taken separately or as a trading system. Performance of trading systems is complicated to evaluate. They need to be programmed in the same exact way, with a general methodology. Its objective is to be able to compare the results and estimate if they issued a significant performance which can cause a bias in efficient market hypothesis.

JEL Classification: G12, G14

[1] INSEEC Business School – Finance Department, 27 Rue Vellefaux, 75010 Paris,
e-mail: jsenanedsch@groupeinseec.com

1 Introduction

The efficient market hypothesis is one of the most important principles of modern finance. It has been introduced by Fama in 1965 and is often the centre of debates during financial crises. Many studies have proven that financial markets are generally very close from being efficient but traders keep trying every day to prove it wrong by using forecasting strategies as technical analysis. The efficient market definition proposed by Jensen [1] is generally the most accepted: "A market is efficient with respect to information set θ_t if it is impossible to make economic profits by trading on the basis of information set θ_t…. By economic profits, we mean the risk-adjusted returns net of all costs."

This definition means that:

- On an efficient market, it would be impossible to consistently achieve returns in excess of the market average return on a risk-adjusted basis.

- Costs are included in the calculation of returns. These costs could be transaction costs, information search costs and active portfolio management costs which include the time spent as well as opportunity costs.

- It is fundamental to calculate the performance of a strategy rather than the strict return. The risk is a major characteristic of an investment strategy.

In the financial literature, the *ex post* Sharpe ratio is the factor most frequently used to determine whether a portfolio generated excess returns. A benchmark portfolio is typically used to compare the Sharpe ratios and the most common strategy for a benchmark portfolio is a simple buy-and-hold policy. An efficient market anomaly is considered to exist if a portfolio consistently or in average outperforms a buy-and-hold policy applied to the same securities as those in the portfolio, after accounting for the costs related to both the portfolio and the

buy-and-hold policy.

Our study investigates strategies resulting from technical analysis. Technical analysis is a set of methods for predicting market movements by using past transaction data to develop buy and sell signals, in order to obtain economic profits (or excess returns) on a risk-adjusted basis. Each of the methods used in technical analysis can be considered a test of market efficiency since it assesses the excess returns obtained in a market over a given period.

This paper contains five sections including introduction and conclusion. Section 2 explains the principle of simulations. Section 3 studies the algorithms used for modelling strategies. Finally, section 4 develops performance and significance measurements.

2   Simulation Principle and Test Procedure

2.1 Simulation and Back Testing

An empirical simulation process applied to technical analysis is the more effective way to test its performance. The objective of a simulation is to reproduce the same signals used by traders. These signals are then applied to historical data in the same realistic conditions. The performance of technical analysis methods is almost always evaluated using simulations. Nevertheless some researchers propose to evaluate the predictive capacity of these methods. It is different from an evaluation of performance and is generally assessed through other means, such as by studying the short and medium term behaviour of prices after signals are issued. However, high predictive capacity does not guarantee good performance.

Due to the nature of this empirical test, some assumptions (some by choice, others are mandatory) have to be made to perform a simulation. First, we have to decide whether or not the model would allow short selling. The best choice is to simulate both cases, with and without short positions. However, it is generally

accepted to extrapolate the performance calculate without short positions to a general model. In this case, a sell signal issued when an investor does not own the corresponding security (because it has not yet been bought or already been sold) is considered as a confirmation of a sale (or of non-ownership) rather than a signal prompting a short sale. This choice is often accepted to be a way of making technical analysis methods perform better than methods that do allow for short positions. Second, a model should specify the number of shares bought or sold for each signal. Some problems are often encountered in this axiom of programming because it can be necessary to define strength for each signal depending on a precise variable. In general, for each buy or sell signal, a model executes a transaction for one share. We can then extrapolate results to higher volumes: one signal would equal X shares. If this choice is made, this model would then treat two consecutive buy signals as a buy signal followed by a hold signal rather than a double buy.

In this context, the strategies applied do not change the historical price data. Indeed, all empirical studies of the performance of technical analysis methods use actual historical data that cannot be modified. Therefore another assumption made is that, given the large transaction volumes on data selected, transactions do not affect stock prices.

Finally, modelling a trading system requires to define the perfect moment when a signal is effective. Actually, data are often collected automatically though databases such as Bloomberg or Reuters after a trading day and the signal can be effective only for the next day at the opening price.

2.2 Methodology

The general methodology can be used as an empirical test of efficient market hypothesis. The steps have to be followed precisely as illustrated in Figure 1:
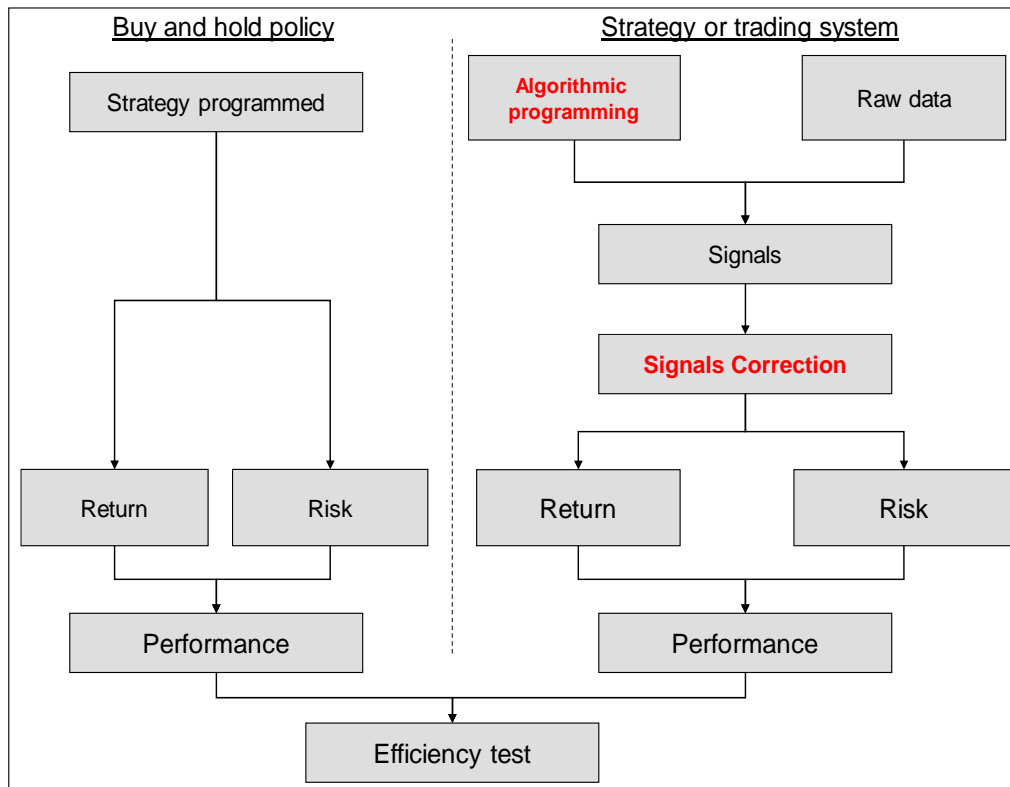
Figure 1:   General test procedure

- Data: the first step of a study is to determine which type of assets to test and select the time period and interval for the corresponding data. For testing technical analysis methodology, academic literature prefers to study individual stocks rather than market indices because indices are subject to a survivorship bias [2]. We recommend using all securities from an index to avoid any data snooping bias. Banking and insurance companies can be kept in the sample. It is also preferable to work with stocks with high liquidity because a simulated strategy does not contribute to change the price proposed in the real data and performance can be affected in certain cases. We should also identify the high and low prices for each time interval; for example, for the weekly data sample, we selected the highest stock price in each week. This gave us a data sample

with the opening, closing, high, and low prices for each time interval. All of the data should be considered as 'out-of-sample' since they are not used to build the technical analysis tools employed in the study.

- Benchmark: the performance of the strategy or trading system must be compared to a benchmark. The absolute value found by the evaluation of the performance has no meaning. The relative value i.e. the comparison between the performance of the strategy and the performance of the benchmark is very useful to know whether or not this strategy should be accepted and if it issues a bias in the efficient market hypothesis. The most common benchmark is called 'buy-and-hold policy' and consists of a naïve long position during the whole period of test.

3   Simulation Process

3.1 Definition of Patterns

To program a strategy or a trading system it is necessary to develop formulas to make a systematic analysis of the signals. Some specialists use figure recognition process to find patterns on the price charts. However, we can recommend the most used method which consists to develop an algorithmic model. Algorithms are very precise and can define any pattern, figure or trading system used in technical analysis. Those algorithms have to be programmed with dynamic variables to be able to change values for adapting a possible learning process. Caginalp and Laurent [3] provide a methodology to create algorithms from a simple definition of technical patterns. In this section, we provide our own algorithms to give an example of definition of the chart in Figure 2.

In order to simulate these patterns and to develop algorithms, the first step is to describe clearly the strategy with rules making the relations between variables. We can first review the literature on the specific strategy if it a common one or a

variant of a common one.

Algorithms have to precise the conditions of emission of a signal and what kind of signals is programmed. In general, each pattern has an inverse, and the pattern and its inverse form a pair—one buy order and one sell order—called basic trading system or strategy.

To illustrate the simulating process using algorithms, we propose an example with Japanese candlesticks, a classical technical analysis method. Let O_n be the opening price on day n; t the last known day (i.e., the day on which the signal was issued that triggered the buy or sell transaction at market open the next trading day); and C_n, H_n, and L_n the closing, high, and low prices on day n, respectively. These are the same variable names as those used by Nison [4]; we also used the same pattern name. Below is a definition of the pattern we propose as example along with the algorithms we created.
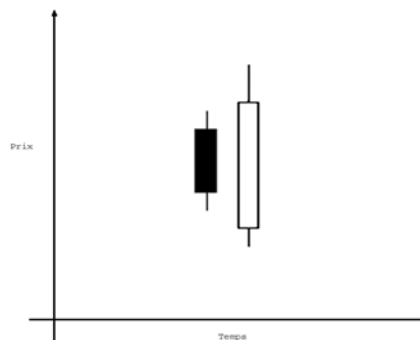

Figure 2:   Bullish Engulfing

Definition 1 *(Engulfing)* – The Bullish Engulfing (*UENGULF*) and Bearish Engulfing (*EENGULF*) patterns consist of a short and long candlestick side by side, with each candlestick a different colour (one black, one white).

$$UENGULF \equiv \begin{cases} C_{t-1} < O_{t-1} \\ C_t > O_t \\ C_t \geq O_{t-1} \\ C_{t-1} \geq O_t \\ C_t - O_t > O_{t-1} - C_{t-1} \end{cases}$$

$$EENGULF \equiv \begin{cases} C_{t-1} > O_{t-1} \\ C_t < O_t \\ O_t \geq C_{t-1} \\ O_{t-1} \geq C_t \\ O_t - C_t > C_{t-1} - O_{t-1} \end{cases}$$

3.2 Signals Precisions and Trading Systems

A trading system is created when a strategy combines a multiple amount of tools to generate signals. By nature, those signals may not be coherent together and need a correction. There are multiple ways of correcting the signal issued from technical analysis. Our study proposes one possibility which can be adapted following the specific cases studies. Table 1 gives precisions to apply these corrections.

First of all, some rules have to be established in order to process the corrections. As an illustration, we can provide a short example:

- A strategy must start with a buy signal.
- A strategy must finish with a sell signal.
- A buy (sell) signal can only be followed by a sell (buy) signal.
- In case of contradictory signals in the same day, no signal is logged.

This methodology can be used, in theory, with any trading system. Some of them will be more complex and some improvements should be made. We tested these principles with a combination of a large number of graphic and systematic strategies with success. The goal is to keep the same logical process and adapt the specifics to the trading system.

Table 1 : Example of signals correction for trading systems

Day	Strategy 1	Strategy 2	Strategy corrected
1	Buy	Buy	Buy
2		Sell	Sell
3	Sell	Buy	
4		Sell	
5	Buy		Buy
6		Buy	
7		Sell	Sell

4   Performance and Significance Measurement

4.1 Performance

We propose to use the *ex post* Sharpe ratio to measure risk-adjusted performance. This method is backed by financial theory and is the most common indicator used by asset management professionals [5]. It is then easier to compare the results with other studies from the literature. Many studies justify the preference for Sharpe ratios. Pedersen and Rudholm-Alfvin [6] evaluate rakings of different securities calculated from eight different performance indicators from 1998 to 2003. After calculating the correlations of these ranks, they found that the Sharpe ratio exceeds 90% in most of the cases. Eling and Schuhmacher [7] and Eling [8] confirm these results for respectively 1994-2003 and 1996-2005.

Sharpe [9] introduced the Sharpe Ratio also called *reward-to-variability ratio*. Let R_{it} be the return of the strategy i in period t and R_f the return of the risk free asset:

$$D_t = R_{it} - R_{ft}$$

Letting $\overline{D}$ be the average value of D_t over the historic period from t=1 through T:

$$\overline{D} = \frac{1}{T}\sum_{t=1}^{T} D_t$$

and σ_D be the level of risk calculated from the standard deviation of D over the period:

$$\sigma_D = \sqrt{\frac{\sum_{t=1}^{T}(D_t - \bar{D})^2}{T-1}}$$

Thus, the ex post Sharpe ratio is:

$$S_h = \frac{\bar{D}}{\sigma_D}$$

As mentioned previously, this ratio should be calculated on the strategy and on the benchmark. We then compare those two ratios to evaluate the real performance of the strategy.

It is best to use this ratio with positive returns. In fact, this ratio is biased when values are negative. For example let A and B be two possibilities of investment, $\bar{D}_{tA} = $ - 5% ; $\sigma_{tA} = 0{,}25$ et $\bar{D}_{tB} = $ - 5% ; $\sigma_{tB} = 0{,}5$. We obtain Sharpe ratios of -0.2 and -0.1 respectively, while we prefer investment A, less risky for the same risk premium.

To evaluate the return of a strategy used as an input in the Sharpe ratio, it is necessary to first simulate the strategy to generate the signals. Then, these signals will allow finding the buying and selling price of the asset. For each buy signal, the strategy is debited with the purchase price plus transaction costs and for each sell signal, the strategy is credited with the sale price minus fees; return comes directly from the sum of these movements from all the signals.

Obviously, results must take into account the adjustment to the opportunity cost. It comes from the systematic investment on risk-free rate for the money not invested. Dividends are also invested at risk free rate when they are paid. Risk-free rate can be chosen from a lot of different rates. Literature prefers in general 10-year government bonds yield from the studied area or the best AAA yield.

A lot of simulations are made without including transaction costs. This

analysis is interesting to know the inherent performance of a strategy but it is necessary to add another analysis including transaction costs to evaluate the performance as mentioned in the definition of Jensen [1]. It is actually quite complex to assess the transaction costs rate to apply in this general concept, it will depend on the market and the assets. Literature often considers that average transaction costs should be equal to 1% of the value of the transaction applied on a sell and on a buy situation[2]. In a more practical way, it is clear that online brokers or institutional traders have access to a lower rate. Corrado and Lee [10], Allen and Karjalainen [11], Neely [12] and, Korczak and Roger [13] among others propose a rate of 0.25% which seems to be more realistic to modern markets.

4.2 Significance

We propose to use Monte Carlo simulations to determine whether results (in terms of performance) are statistically significant. Lo, Mamaysky and Wang [14] and Ready [15] have concluded that simulations with geometric Brownian motions are the most appropriate to use in significance tests when assessing the predictive power and performance of technical analysis methods. The Student test assumes that returns are normal and stationary, and the bootstrap can lead to autocorrelations and other biases that do not appear in Monte Carlo simulations. The Monte Carlo simulations applied to technical analysis come directly from the work done by Lo, Mamaysky, and Wang [14], and incorporates the findings of Efron [15] and Marshall, Young, and Rose [16] for generating high and low prices.

---

[2] Articles applying this rate: Cootner [17], Alexander [18], Levy [19] [20], Van Horne and Parker [21], Akemann and Keller [22], Solt and Swanson [23], Brush and Boles [24], Fyfe, Marney and Tarbert [25], Parisi and Vasquez [26].

Monte Carlo simulations should, at least, generate 100 series of artificial days for each stock, and then use detection algorithms on these simulated series. For each stock, artificial series of high, low, opening, and closing prices should be generally simulated. In this model, the series are governed by a geometric Brownian motion and have the same two first motions as those in the base series, as well as the same sample size and initial value.

This significance test is based on a key difference between the simulated series and the base series. When technical analysis patterns are used with actual data, the underlying assumption is that the patterns have a predictive power. However, when the patterns are used with a simulated series they have no predictive power, because by definition it is impossible to predict the outcome of a geometric Brownian motion. Therefore any excess returns generated by the technical analysis methods are strictly random and substantiate the efficient market hypothesis.

We test the null hypothesis under which the prices in the original series follow a geometric Brownian motion, meaning that no excess returns are generated on the base price series. Therefore if technical analysis patterns genuinely existed in the original series, one would see a significant difference between the returns generated on the base series and those on the simulated series. The acceptance probability for the null hypothesis is the number of times (in percentage) that the returns from the simulated series exceed those from the base series.

4.3 Sample and data-snooping

Data sample used to evaluate the performance of a strategy should be large enough to be considered significant. In general, the equivalent of three years of daily data is the minimum to be able to run a performance test with significant results. Data sample should also cover an entire economic cycle with periods of

rising and falling of prices to prove the systematic results of the test. It is possible to test a strategy applied for example on daily, weekly and monthly data or even high frequency data so as to analyse several time intervals and test each strategy thoroughly based on multiple results. This procedure leads to a problem of data-snooping. Actually, data should not be chosen ex-post because it would bring a subjective aspect in the test. In order to prevent a data-snooping bias, we propose to divide the data into sub-periods. For example, we can first divide the total data sample into three of four sub-periods then create a sub-period starting six months after the first day of the overall data period.

One should be careful to the link between data and the kind of strategy simulated. Actually, if we use relatively old data, a study can be subject to Cooper and Gulen's [27] conclusion which considers that a valid efficiency test cannot be carried out with a model developed from software that did not exist during the time period over which the data for the test are taken.

5   Conclusion

The objective of our study is to facilitate programming and testing of strategies on financial markets via our general methodology. This approach includes an understanding of several concepts: technical analysis, efficient market hypothesis, simulation, benchmark, signals correction, algorithmic models and performance with significance measurement. Our general methodology allows researchers to test efficient market hypothesis and helps specialists to obtain the real non-biased performance of their strategies.

References

[1]   M.C. Jensen, Some Anomalous Evidence Regarding Market Efficiency, *Journal of Financial Economics*, **6**,(1978), 95-101.

[2]   R. Sullivan, A. Timmermann and H. White, Data Snooping, Technical Trading Rule Performance and the Bootstrap, *Journal of Finance*, **54**, (1999), 1647-1691.

[3]   G. Caginalp and H. Laurent, The Predictive Power of Price Patterns, *Applied Mathematical Finance*, **5**, (1998), 181-205.

[4]   S. Nison, *Japanese Candlestick Charting Techniques*, New York Institute of Finance, New York (1991).

[5]   A. Lo, Bubble, rubble, finance in trouble? *Journal of Psychology and Financial Markets*, **3**, (2002), 76-86.

[6]   C.S. Pedersen and T. Rudholm-Alfvin, Selecting risk-adjusted shareholder performance measure, *Journal of Asset Management*, **4**(3), (2003), 152-172.

[7]   M. Eling and F. Schuhmacher. Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking* and Finance, **31**(9), (2007), 2632-2647.

[8]   M. Eling, Does the measure matter in the mutual fund industry?, *Financial Analysts Journal*, **64**, (2008), 1-13.

[9]   W. Sharpe, Mutual Fund Performance, *Journal of Business,* **39**(1), (1966), 119-138.

[10] C. J. Corrado and S.H Lee, Filter Rule Tests of the Economic Significance of Serial Dependencies in Daily Stock Returns, *Journal of Financial Research*, **15**, (1992), 369-387.

[11] F. Allen and F. Karjalainen, Using Genetic Algorithms to Find Technical Trading Rules, *Journal of Financial Economics*, **51**, (1999), 245-271.

[12] C. Neely , Risk-Adjusted, Ex Ante, Optimal Technical Trading Rules in Equity Markets, *International Review of Economics and Finance*, **12**, (2003), 69-87.

[13] J. Korczak and P. Roger, Stock Timing Using Genetic Algorithms, *Applied Stochastic Models in Business and Industry,* **18**,(2002),121-134.

[14] A. Lo, H. Mamaysky and J. Wang, Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation, *Journal of Finance*, **56**, (2000), 1705-1765.

[15] M. J. Ready, Profits from Technical Trading Rules, *Financial Management,* **31**, (2002), 43-61.

[16] B. Efron, Bootstrap methods: Another look at the jackknife, *The Annals of Statistics*, **71**, (1979), 1-26.

[17] B. Marshall, M. Young and L. Rose, Market Timing With Candlestick Technical Analysis, *Journal of Financial Transformation, Securities Issue*, (2007), 18-25.

[18] P. Cootner, Stock Prices: Random vs. Systematic Changes, *Industrial Management Review*, **3**, (1962), 24-45.

[19] S. Alexander, Price Movements in Speculative Markets: Trends or Random Walks No. 2, *Industrial Management Review,* **5**, (1964), 25-46.

[20] R. Levy, The Principle of Portfolio Upgrading, *Industrial Management Review*, (1967a), 82-96.

[21] R. Levy, Random Walks: Reality or Myth, *Financial Analysts Journal*, **23**(6), (1967b) 69-77.

[22] J. Van Horne and G. Parker, Technical Trading Rules: A Comment, *Financial Analysts Journal*, **24**, (1968), 128-132.

[23] C.A Akemann and W.E Keller, Relative Strength Does Persist, *Journal of Portfolio Management*, (1977), 38-45.

[24] M. Solt and P. Swanson, On the Efficiency of the Markets for Gold and Silver, *Journal of Business*, **54**(3), (1981), 543-478.

[25] J.S. Brush and K.E. Boles, The Predictive Power in Relative Strength and CAPM, *Journal of Portfolio Management*, (1983), 20-23.

[26] C. Fyfe, J. Marney and H. Tarbert, Technical trading versus market efficiency - a genetic programming approach, *Applied Financial Economics*, **9**, (1999), 183-191.

[27] F. Parisi and A. Vasquez, Simple Technical Trading Rules of Stock Returns: Evidence from 1987 to 1998 in Chile, *Emerging Markets Review*, **1**, (2000), 152-164.

[28] M. Cooper and H. Gulen, Is Time-Series Based Predictability Evident in Real-Time?, *Working Paper*, *Krannert Graduate School of Management, Purdue University*, (2003).