Candlestick technical trading strategies: Can they create value for investors?

Ben R. Marshall a,*, Martin R. Young a, Lawrence C. Rose b

a Department of Finance, Banking and Property, College of Business, Massey University, Private Bag 11222, Palmerston North, New Zealand
b Department of Commerce, College of Business, Massey University, Private Bag 102 904, North Shore Mail Centre, Auckland, New Zealand

Received 17 June 2005; accepted 1 August 2005
Available online 28 November 2005

Abstract

We conduct the first robust study of the oldest known form of technical analysis, candlestick charting. Candlestick technical analysis is a short-term timing technique that generates signals based on the relationship between open, high, low, and close prices. Using an extension of the bootstrap methodology, which allows for the generation of random open, high, low and close prices, we find that candlestick trading strategies do not have value for Dow Jones Industrial Average (DJIA) stocks. This is further evidence that this market is informationally efficient.

© 2005 Elsevier B.V. All rights reserved.

JEL classification: G12; G14

Keywords: Candlestick; Technical analysis; Bootstrapping; Trading strategies

1. Introduction

Technical analysis uses past prices and other trade statistics, such as volume, to predict future prices movements. Proponents of technical analysis believe that these data contain important information about changes in investor sentiment and that the reaction to news
is a gradual process which allows trends to develop. Despite its widespread acceptance and adoption by practitioners, technical analysis has been described by Malkiel (1981) as an “anathema to the academic world”. This is due to its conflict with market efficiency, one of the central pillars of academic finance. In an efficient market, prices are said to reflect information to the point where the potential risk-adjusted excess returns of acting on information do not exceed the cost of doing so (Jensen, 1978). Any information contained in past prices is therefore reflected in current prices, making the study of past prices a futile exercise.

Early academic studies of technical analysis by Alexander (1964) and Fama and Blume (1966) find that technical analysis’ profitability is eroded by transaction costs. This finding, which is consistent with market efficiency, resulted in few further technical analysis studies over the next twenty years. More recently, the increasing popularity of technical analysis with practitioners and the growing evidence that investors do not always act rationally, have caused academic finance to take another look at technical analysis.¹

This paper considers the profitability of short-term technical analysis by focusing on candlestick technical trading strategies. Candlestick charting, which was originally applied to rice markets in the 1700s, is designed to capture short-term price movements and is therefore most useful for horizons of approximately ten days (Nison, 1991). Since its introduction to the Western World candlestick technical analysis “have become ubiquitous, available in almost every software and online charting package” (Nison, 2004, p. 22).

To the best of our knowledge, this is the first paper which rigorously and independently analyses these strategies. Further, unlike previous technical analysis literature on stocks, which has mostly used data from an index such as the Dow Jones Industrial Index (DJIA), this study uses data for the individual stocks that comprise the DJIA for the period 1 January 1992 to 31 December 2002.

Candlestick technical analysis involves the relationship between open, high, low and close prices. These four prices are displayed as objects that resemble candles as shown in Fig. 1. When the close is above (below) the open the candle “body” is white (black).

Few empirical studies are free from data-instigated pretest biases. Lo and MacKinlay (1990) and Lakonishok and Smidt (1988) point out that we can expect the degree of such biases to increase with the number of published studies performed on any single data set and propose that new data is the best protection against data snooping bias. Sullivan, Timmermann and White (STW) (1999) suggest that data snooping should be accounted for in technical trading rule tests by considering the performance of the best trading rule in the context of the full universe of trading rules from which the rule is conceivably chosen. However, LeBaron (2000) and Ready (2002) point out that the STW (1999) data snooping adjustment technique is not perfect, as it depends on simulating the snooping process that has been occurring. There are no formal tests to ascertain this.

Candlestick technical analysis is more robust to the criticism of data snooping than are tests of other technical trading rules such as the moving average, as it was developed for an entirely different purpose, forecasting rice markets. Testing candlestick technical analysis using US stock data is therefore, most clearly, an out-of-sample test. This approach even

¹ Technical analysis methods tested include moving average trading range break out rules (Brock et al., 1992), genetic algorithms (Allen and Karjalainen, 1999), chart patterns (Lo et al., 2000), and Dow Theory (Brown et al., 1998). More recently, studies have found that technical analysis is linked to order placement patterns (Osler, 2003; Kavajecz and Odders-White, 2004).
survives the weak criticism that it is simply a test of another technical trading rule on US data. The use of open, high, low and close prices by candlestick technical analysis differentiates it from previous technical trading studies that have used close price data only.

Data choice is critically important to studies of technical analysis for reasons other than data snooping. We suggest that this paper’s use of DJIA component stock data for the 1992–2002 period has several advantages over the more traditional choice of 50–100 years of DJIA data. Firstly, until the recent introduction of the Diamonds Exchange Traded Fund, the DJIA was not able to be traded in its own right. Any technical trading signals on the DJIA would therefore be unable to be implemented without purchasing each of the DJIA components in the correct proportions. Secondly, as Day and Wang (2002) document, tests of technical trading rules on index data can be biased due to nonsynchronous trading. Thirdly, market efficiency claims only that prices reflect all known information at that point in time, not information that may come to light in the future. Recently developed technical trading rules that are reliant on substantial computer power, which reveal profits on 50–100 years of historical data, are therefore not necessarily evidence against market efficiency (Miller, 1990). For this reason the start point of 1 January 1992 is carefully chosen. Despite being a popular trading technique in Japanese financial markets for some considerable time, the seminal candlestick trading strategy book in English was not published until 1991.2 Major data providers, such as Reuters, also started making open, high, low and close data available from the middle of 1991. Therefore users of technical analysis would have been aware of candlestick techniques and have had the ability to implement them only from the start of 1992.

Finally, technical analysts claim that technical analysis is most reliable for actively traded stocks (Morris, 1995). This makes the DJIA component stocks an obvious choice. They are also an important choice from a market microstructure perspective. Trading on the NYSE begins with a call auction. At the open, the specialist sets a single price at which the accumulated order imbalance from market-open and limit orders clears (Madhavan and Panchapagesan, 2000). The assumption that investors could buy DJIA component stocks at the recorded opening price therefore seems reasonable.

The profitability of candlestick trading strategies is tested using a bootstrap methodology. An extension of the methodology developed by Efron (1979) for one series is used to generate random open, high, low and close series. This methodological advance can be applied to other trading rules that rely on signals from closely related series. The profits

---

accruing to technical trading strategies on the actual DJIA component stock series are then compared to the profits earned on the random series to determine whether or not the profitability of technical trading strategies is statistically or economically significant.

We find strong evidence that candlestick technical analysis is not profitable on the US stock market. Neither bullish nor bearish signals consistently outperform a buy-and-hold approach. This is further evidence that the US stock market is informationally efficient.

The remainder of the paper is organized as follows: Section 2 describes the data source and candlestick patterns. Section 3 outlines the approaches used to measure candlestick trading strategy profitability. Section 4 includes a discussion of the empirical results and Section 5 considers their robustness. Section 6 concludes the paper.

2. Data and candlestick patterns

2.1. Data

Price data in open, high, low and close format are sourced from Reuters. These data are not adjusted for cash dividends. Many studies of technical analysis ignore dividends due to their focus on index data and the difficulty associated with adjusting an index for dividends. However, Day and Wang (2002) point out that excluding dividends biases the buy and hold return downwards, and favours technical analysis. They therefore recommend the inclusion of dividend data. Following Day and Wang (2002) we add cumulative dividends on to each of the four price series for each stock at each ex-date. Dividend data are sourced from CRSP.

The sample includes stocks that are part of the DJIA index for the 1 January 1992–31 December 2002 period. The starting point is carefully chosen to ensure that investors would have been aware of candlestick technical analysis and have had the ability to apply it. These two factors are required for any test of market efficiency. Technical analysis is said to be most effective on actively traded stocks. For this reason data for the period that a stock is actually in the DJIA are used. When a stock is replaced in the DJIA it is replaced in this study with its replacement in the DJIA (with three exceptions). During the period of the study there were eight changes made to the DJIA. Reuters data were missing for three companies, Westinghouse Electric, Texaco Incorporated and Woolworth. These were replaced in the DJIA on 17 March 1997 by Travelers Group (now Citigroup), Hewlett-Packard Company and Wal-Mart Stores, respectively. Each of these replacement companies was actively traded prior to its inclusion in the DJIA so all are included in the sample for the entire period.

2.2. Candlestick patterns

Candlestick technical analysis is credited to Munelusa Homma, a legendary rice trader who amassed a huge fortune through applying these rules to his local rice exchange in Osaka. Candlestick technical analysis was undiscovered in the western world until 1991 when Steve Nison published “Japanese Candlestick Charting Techniques: A Contemporary Guide to the Ancient Investment Techniques of the Far East” (Pring, 2002).

---

3 These data are adjusted for stock splits and stock dividends.
A daily candlestick is a graphical representation of the day’s open, high, low and close prices. Daily candlesticks are commonly referred to as “single lines”. Some single lines are said to have forecasting power in their own right. For instance, a White Marubozu (shown in Fig. 2) is said to be a single line that suggests further price increases because prices open at the day’s low and rise throughout the day to close at the day’s high. A White Marubozu is said to indicate a situation where buyers overwhelm sellers and bid up prices during the day. The odds are that this supply/demand imbalance will lead to further price rises in the future. Other single lines are neutral giving no indication of future price movements.

Together, consecutive single lines can form continuation and reversal patterns. Continuation patterns indicate that the prevailing trend will continue, while reversal patterns suggest that there will be a change in trend. All single lines and most continuation and reversal patterns have a bullish and a bearish variety. In this context, the term bullish (bearish) suggests future price increases (decreases).

There are numerous combinations of single lines that are neither continuation nor reversal patterns. In addition, some continuation and reversal patterns are said to have very little, or no, forecasting power. To determine whether a continuation or reversal pattern has strong forecasting power, proponents of candlestick technical analysis developed a system of combining the two or three individual single lines that make up the pattern to form an overall single line for the two- or three-day period. The characteristics of this overall single line are supposed to indicate whether or not the pattern does have forecasting power.

The rule for combining the single lines that make up a pattern into an overall single line is as follows: the combined high is the high on individual single lines, the combined low is the low on individual single lines, the combined open is the open from the first single line, and the combined close is the close from the last single line (Morris, 1995).

An example of a bullish reversal pattern is the Bullish Engulfing pattern (shown in Fig. 3). The Bullish Engulfing pattern involves a short black candle being followed by a long white candle which opens below, but closes above, the previous day. The overall single line formed by combining the two individual single lines that make up the Bullish Engulfing pattern is bullish, which confirms that the Bullish Engulfing pattern is said to have power to predict price increases.

In selecting the single lines and patterns to test we have adopted the following approach. Firstly, we identify all single lines (18) and patterns (44) documented by practitioner books. The material in these books is checked against an English translation of Shimizu (1986), the seminal Candlestick book in Japanese, to ensure that nothing from the Japanese candlestick literature is missing from, or adapted by, these books. Secondly,
we identify and remove all single lines (4) and patterns (22) that are supposed to have zero explanatory power. We use the method of forming an overall single line from a pattern, as documented by Morris (1995) and Nison (1991). Finally, we identify single lines (0) and patterns (8) that occur very infrequently. These are defined as those that occur fewer than 10 times in our total sample. This leaves us with 14 single lines and 14 reversal patterns to test. A detailed description of the candlestick single lines and patterns we test is provided in Appendix A.

Although the universe of candlestick single lines and patterns is greater than those tested in this paper, this approach results in tests of single lines and patterns that are most likely to be used by exponents of candlestick technical analysis. They are certainly the ones that the candlestick technical analysis literature says have explanatory power. It therefore seems logical to use these rules. As a parallel for this approach, Brown et al. (1998) consider the power of the Dow Theory by testing returns on days following Dow buy signals, not the returns on days that the Dow Theory did not signal a buy.

We define single lines and patterns as they are outlined in the major candlestick technical analysis books. Published books are explicit on some issues. For example, when a white single line is required to have similar open and low prices, Morris (1995, p. 25) states that the difference between these two prices “should be less than 10% of the open–close range”. However, candlestick books point out that there is some flexibility in defining other aspects of single lines such as the distance between open and close for the candle to be classified as a long candlestick.

Single lines are said to have forecasting power regardless of the underlying trend in the market. In contrast, reversal patterns require the existing trend to be identified. Candlestick technical analysis is a short-term technique so candlestick books advocate that a ten-day moving average of prices is used to determine the trend. If price is above (below) the ten-day moving average an uptrend (downtrend) is said to exist (Morris, 1995). Following Morris (1995), our base tests use an exponential moving average which gives more weight to the most recent observations.

The challenge of correctly specifying technical trading rules is faced by all researchers in this area. In fact, the issue is more serious in papers such as Lo et al. (2000) (hereafter, LMW) which test patterns, such as the head and shoulders formation, that are far more difficult to define. LMW (2000, p. 1714) state that they “settle on an acceptable bandwidth for their pattern detection algorithm by trial and error”. In this research, this issue affects only single lines as published books are clear on what combinations of single lines constitute a continuation or reversal pattern. As a final check of our single line specifications, we conduct sensitivity analysis to see how changing the single line and trend definition affects the results in terms of both the number of patterns and profitability.
3. Measures of candlestick trading strategy profitability

The profitability of candlestick trading strategies is tested using a bootstrapping methodology (e.g., Efron, 1979). Morris (1995, p. 213) points out that “candlestick analysis is short-term. Any patterns that give longer-term results are surely just coincidental”. Morris (1995) defines the maximum period that candlestick technical analysis has value as ten days. We conduct sensitivity analysis around the holding period, but our core tests are for ten days. The methodology description is therefore based on a ten-day holding period.

Our approach is to first investigate whether there is any statistical significance to the profits from following candlestick signals. Consistent with Brock et al. (1992) (hereafter, BLL) we use raw returns rather than excess returns. This approach is appropriate for short-term rules, as variations in the risk premia are likely to have a long periodicity relative to the holding period (Sweeney, 1986).

The bootstrapping methodology has several advantages over the more traditional t-test. Firstly, unlike t-statistics, bootstrapping can accommodate well known characteristics of stock return data such as skewness, leptokurtosis (fat tails), autocorrelation, and conditional heteroskedasticity. A second benefit of the bootstrap methodology is that it allows us to simulate distributions of the trading rule returns by any specified model.

The first step in applying the bootstrapping methodology is the choice of null models to fit the data. To ensure consistency with previous papers, we adopt four widely used processes for stock prices: a random walk, an autoregressive process of order one (AR(1)), a GARCH in-Mean (GARCH-M) model and an Exponential GARCH (EGARCH) model. The parameters of the later three models are estimated separately for each stock series for the entire eleven years. The results are very similar across all null models so we focus on the results from the GARCH-M model, the most common null model in the literature.

Previous papers have all tested trading rules that are based solely on close prices. Although we are considering open, high, low, and close prices we start by replicating previous approaches. This involves fitting the GARCH-M null model to the original close price series. This process is conducted separately for each stock because it does not make sense to try and fit a null model to a long series of returns that has been created by joining together series of individual stock returns.

The GARCH-M model is shown below in Eqs. (1a)–(1c)

\[ r_t = \alpha + \gamma \sigma^2_t + \beta \varepsilon_{t-1} + \varepsilon_t, \]  
\[ \sigma^2_t = \sigma^2_0 + \alpha \varepsilon^2_{t-1} + \beta \sigma^2_{t-1}, \]  
\[ \varepsilon_t = \sigma_t z_t, \quad z_t \sim N(0, 1). \]

In this model, the error, \( \varepsilon_t \), is conditionally normally distributed and serially uncorrelated. The conditional variance, \( \sigma^2_t \), is a linear function of the square of the last period’s errors and of the last period’s conditional variance, which implies positive serial correlation in the conditional second moment of the return process. Periods of high (or low) volatility are likely to be followed by periods of high (low) volatility. The conditional returns in this model are a linear function of the conditional variance and the past disturbance, \( \varepsilon_{t-1} \). Under this return-generating process, volatility can change over time and the expected returns are a function of volatility as well as of past returns.\(^5\) The parameters

\(^5\) See Engle et al. (1987).
and standardized residuals are estimated for each DJIA component stock using the maximum likelihood criterion.

In accordance with BLL (1992), we standardize the residuals of our GARCH-M model using estimated standard deviations for the error process. The standardized residuals are then redrawn with replacement to form a scrambled residuals series which is used, along with the estimated parameters, to form new representative close return series. These returns are then exponentiated to form new close price series for each stock. These scrambled series have the same drift in prices, the same volatility, and the same unconditional distribution. However, by construction the returns are independent and identically distributed. The standardised residuals are not restricted to a particular distribution, such as Gaussian, by this procedure.

Once a randomly generated close series has been formed, we create vectors of the original (high-close)/close and (close-low)/close percentage differences. We then take a random sample from these percentage difference vectors. Next, we add (subtract) these high-close (close-low) percentage differences to (from) the simulated close price to form simulated high and low prices. We use a similar process to generate simulated open prices. To ensure that the resampled open price is never higher than the high nor lower than the low, we resample the close–open percentage differences if this situation arises.

This process is replicated 500 times for each stock so we end up with 500 simulated sets of open, high, low and close series for each stock in our sample for each null model. Efron and Tibshirani (1986) suggested that 500–1000 simulations are enough to approximate the true estimator. We also find convergence before 500 simulations so do not proceed with additional simulations.

The proportion of times that a candlestick trading rule produces more profit on the bootstrapped series than on the original series following a signal is a simulated p-value for the null hypothesis that the trading rule has no value. For a candlestick to have statistically significant forecasting power at the 5% level, the simulated p-value should be less than 0.05. In other words, more profit should be produced on the random series than the original less than 5% of the time. In order to provide an indication of the overall statistical significance of candlestick technical analysis on DJIA stocks, we report a count of the number of DJIA stocks (out of 35) for which the profits following a candlestick signal are statistically significantly greater (at the 5% level) than buy-and-hold profits.

To check the robustness of our results, we investigate the variation in profits stemming from entering the market following a signal at close \( t \), close \( t + 1 \), and open \( t + 1 \), where \( t \) is the day that the signal is received. When we consider entering at the close price, we conduct the bootstrap process as described above and compare the conditional returns on the bootstrapped close series versus the original close series. When we consider entering at the open price and therefore compare the open returns on the bootstrapped series to those on the original, we begin by bootstrapping the open series and generate high, low and close series from this in a similar fashion to that outlined for the close price series. In each of the tests, we consider each candlestick pattern in isolation. More specifically, we buy (sell) following a pattern and maintain the position for the holding period. We do not close the position if another candlestick pattern gives an offsetting signal. This approach is consistent with the extant technical analysis papers that consider multiple rules (e.g., BLL, 1992).
4. Empirical results

This section contains the summary statistics for the thirty-five stocks in our sample, profitability statistics for each of the candlestick rules, and bootstrap results which document the statistical significance (or otherwise) of the candlestick rules. The results consistently show that candlestick technical analysis has no value.

4.1. Summary statistics

The summary statistics for the thirty-five stocks that are part of our sample for the period of our study (1 January 1992–31 December 2002) are included in Table 1. There are 83,220 daily returns across the stocks in our sample. Return is defined as the natural logarithm of value relatives. Following LMW (2000) we calculate the mean, standard deviation, skewness and kurtosis of the returns of all the stocks in our sample together. As expected, the mean returns of each of the four series are similar. Volatility is also similar across the four series with high and low only slightly less volatile than open and close. All four series display negative skewness. The four series are all leptokurtic, with high and low displaying this characteristic more strongly than open and close.

We conduct sensitivity analysis (see Section 5) around entering the market following a signal at close and open prices, but our core bootstrap results are based around entering the market at the open price on the day after the signal is generated. This appears to be a more realistic assumption than other papers which assume that a technical trader could buy a stock at the close price on the same day that a signal is generated. In reality, this is very difficult as the close price of the stock is what determines whether a trading signal will be generated. A technical analyst following this approach would have to firstly feed estimates of the close price into his/her trading system to see if they generated a signal. If one did s/he would then need to submit a “market at close” order. At this point one could not be sure that the actual close price would be sufficiently similar to the estimated close price to have generated the signal so there is a risk of acting on an invalid signal.

Another option is entering at the close on the day after a signal. This is certainly achievable, but we propose that it is more likely that a trader would enter the market at the first available opportunity following a technical signal. Hence, the use of the open price from the day following the signal is our preferred baseline.

<table>
<thead>
<tr>
<th></th>
<th>Open</th>
<th>High</th>
<th>Low</th>
<th>Close</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>83,220</td>
<td>83,220</td>
<td>83,220</td>
<td>83,220</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
</tr>
<tr>
<td><strong>Std. Dev.</strong></td>
<td>0.0200</td>
<td>0.0174</td>
<td>0.0187</td>
<td>0.0198</td>
</tr>
<tr>
<td><strong>Skewness</strong></td>
<td>$-0.3710^{**}$</td>
<td>$-0.2376^{**}$</td>
<td>$-0.9980^{**}$</td>
<td>$-0.3939^{**}$</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>37.7218$^{**}$</td>
<td>56.8085$^{**}$</td>
<td>57.0827$^{**}$</td>
<td>36.1342$^{**}$</td>
</tr>
</tbody>
</table>

Results are presented for the full sample of thirty-five DJIA component stocks over the ten-year period. Returns are measured as the log difference of the level of each DJIA stock. The mean, standard deviation, skewness and kurtosis are calculated of the entire return series of all stocks.

** Indicates statistical significance at the 1% level.
4.2. Candlestick statistics

The results presented in this section suggest that candlestick technical analysis is not profitable. Contrary to expectations, the proportion of positive profits following a bullish signal is generally less than 50%. As expected, the proportion of positive profits following bearish candlesticks is usually greater than 50%. However, the mean daily profits following these signals are typically negative, which suggests that some big losses affect the overall profitability of these rules.

Table 2
Candlestick statistics

<table>
<thead>
<tr>
<th>Candlestick</th>
<th>N(Sig)</th>
<th>Profit &gt; 0</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Bullish single lines</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long White</td>
<td>2947</td>
<td>0.4760</td>
<td>0.0000</td>
</tr>
<tr>
<td>White Marubozu</td>
<td>642</td>
<td>0.4581</td>
<td>0.0004</td>
</tr>
<tr>
<td>Closing White Marubozu</td>
<td>1565</td>
<td>0.4726</td>
<td>0.0002</td>
</tr>
<tr>
<td>Opening White Marubozu</td>
<td>1611</td>
<td>0.4703</td>
<td>0.0000</td>
</tr>
<tr>
<td>Dragonfly Doji</td>
<td>270</td>
<td>0.4419</td>
<td>−0.0003</td>
</tr>
<tr>
<td>White Paper Umbrella</td>
<td>567</td>
<td>0.4771</td>
<td>0.0005</td>
</tr>
<tr>
<td>Black Paper Umbrella</td>
<td>727</td>
<td>0.4670</td>
<td>0.0002</td>
</tr>
<tr>
<td><strong>Panel B: Bullish reversal patterns</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hammer</td>
<td>57</td>
<td>0.4965</td>
<td>0.0007</td>
</tr>
<tr>
<td>Bullish Engulfing</td>
<td>252</td>
<td>0.4905</td>
<td>0.0004</td>
</tr>
<tr>
<td>Piercing Line</td>
<td>138</td>
<td>0.4717</td>
<td>−0.0004</td>
</tr>
<tr>
<td>Bullish Harami</td>
<td>115</td>
<td>0.5087</td>
<td>0.0006</td>
</tr>
<tr>
<td>Three Inside Up</td>
<td>17</td>
<td>0.5000</td>
<td>0.0010</td>
</tr>
<tr>
<td>Three Outside Up</td>
<td>56</td>
<td>0.4839</td>
<td>−0.0002</td>
</tr>
<tr>
<td>Tweezer Bottom</td>
<td>354</td>
<td>0.4768</td>
<td>0.0001</td>
</tr>
<tr>
<td><strong>Panel C: Bearish single lines</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long Black</td>
<td>2661</td>
<td>0.5081</td>
<td>−0.0007</td>
</tr>
<tr>
<td>Black Marubozu</td>
<td>557</td>
<td>0.5133</td>
<td>−0.0011</td>
</tr>
<tr>
<td>Closing Black Marubozu</td>
<td>1022</td>
<td>0.5123</td>
<td>−0.0009</td>
</tr>
<tr>
<td>Opening Black Marubozu</td>
<td>1737</td>
<td>0.5144</td>
<td>−0.0005</td>
</tr>
<tr>
<td>Gravestone Doji</td>
<td>191</td>
<td>0.5356</td>
<td>−0.0009</td>
</tr>
<tr>
<td>White Shooting Star</td>
<td>520</td>
<td>0.5171</td>
<td>−0.0005</td>
</tr>
<tr>
<td>Black Shooting Star</td>
<td>465</td>
<td>0.5116</td>
<td>−0.0005</td>
</tr>
<tr>
<td><strong>Panel D: Bearish reversal patterns</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hanging Man</td>
<td>84</td>
<td>0.5167</td>
<td>−0.0009</td>
</tr>
<tr>
<td>Bearish Engulfing</td>
<td>289</td>
<td>0.5000</td>
<td>−0.0007</td>
</tr>
<tr>
<td>Dark Cloud Cover</td>
<td>117</td>
<td>0.5282</td>
<td>−0.0004</td>
</tr>
<tr>
<td>Bearish Harami</td>
<td>396</td>
<td>0.5247</td>
<td>−0.0002</td>
</tr>
<tr>
<td>Three Inside Down</td>
<td>34</td>
<td>0.5265</td>
<td>0.0007</td>
</tr>
<tr>
<td>Three Outside Down</td>
<td>36</td>
<td>0.4889</td>
<td>−0.0002</td>
</tr>
<tr>
<td>Tweezer Top</td>
<td>407</td>
<td>0.5263</td>
<td>−0.0007</td>
</tr>
</tbody>
</table>

Column 2 shows the number of signals emitted by the candlesticks (N(Sig)). Profit > 0 is the proportion of times that profits for a ten-day period following a signal are greater than zero. Mean is the average daily profit for the ten-day holding period following a signal. Profits are calculated based on the assumption that trades are entered at the open price on the day following a signal, that positions are held for ten days, and that a ten-day exponential moving average is used to determine the prior trend.
Results from the bullish single lines and patterns are presented in Panels A and B of Table 2. Profits are based on the assumption that trades are entered at the open price on the day following a signal, that positions are held for ten days, and a ten-day exponential moving average is used to determine the prior trend. We report daily profits, regardless of holding period, to aid comparability across the different holding periods adopted in the sensitivity analysis. \( N(Sig) \) is the number of signals in the data. These range from 17 for the relatively rare Three Inside Up pattern to 2947 for the commonly observed Long White single line. The baseline tests use a ten-day holding period so the number of signals needs to be multiplied by ten to arrive at the number of daily returns used in the statistical tests.

The column \( \text{Profit} > 0 \) reports the proportion of profits following a buy signal that are greater than zero. The profits following all the bullish single lines are greater than zero less than 50% of the time. While this is indicative of a poorly performing rule, it is not definitive as it does not take the size of profits into account. It is possible that a rule that is correct less than 50% of the time yields substantially bigger profits than losses making it profitable overall. In addition, it makes no comparison to unconditional profits. The only bullish reversal patterns to yield returns greater than zero more than 50% of the time are the Bullish Harami and Three Inside Up patterns.

The mean profits conditional on bullish single line and reversal pattern signals are all positive with the exception of the Dragonfly Doji, Piercing Line, and Three Outside Up. This suggests that, consistent with candlestick theory, bullish single lines and reversal patterns generally signal positive future returns. The statistical significance of these profits is considered in Section 4.3.

The results from bearish single lines and patterns are presented in Panels C and D of Table 2. The number of bearish single lines and patterns is similar to the number of their bullish counterparts. Profits are positive over 50% of the time, consistent with ones expectations for bearish candlesticks, with the exception of the Bearish Engulfing and Three Outside Down patterns. Other than the Three Inside Down pattern, the mean profits following the bearish single lines and reversal patterns are all negative. These two results indicate that prices fall following bearish candlesticks over half the time, but when prices rise they increase by a larger amount on average which leads to negative profitability overall.

4.3. Bootstrap results

In general, the bootstrap results support the conclusion that candlestick single lines and patterns have no statistically significant explanatory power for DJIA stocks over the 1992–2002 period. Table 3 contains counts of the number of statistically significant individual stock \( p \)-values together with mean values for the bootstrapped and original series for our base scenario of trades being entered at the open price on the day following a signal, positions being held for ten days, and a ten-day exponential moving average being used to determine the prior trend.

The number in the BS column is the sum (across all DJIA stocks) of the mean number of signals per bootstrap series for each individual DJIA stock. The number in the Dow column is the sum of the number of signals on each individual DJIA stock. These results suggest that the number of signals on each original Dow stock series is broadly consistent with the average number of signals per bootstrap series for each stock.

The \( p \)-value count column is the number of stocks for which the mean profit is statistically significantly greater (at the 5% level) on the original series than the 500
Table 3
Full GARCH-M bootstrap results

<table>
<thead>
<tr>
<th>Candlestick</th>
<th>Number of signals</th>
<th>p-value count</th>
<th>Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BS</td>
<td>Dow</td>
<td>Buoy</td>
</tr>
<tr>
<td>Long White</td>
<td>2961</td>
<td>2947</td>
<td>0</td>
</tr>
<tr>
<td>White Marubozu</td>
<td>644</td>
<td>642</td>
<td>0</td>
</tr>
<tr>
<td>Closing White Marubozu</td>
<td>1369</td>
<td>1565</td>
<td>0</td>
</tr>
<tr>
<td>Opening White Marubozu</td>
<td>1615</td>
<td>1611</td>
<td>0</td>
</tr>
<tr>
<td>Dragonfly Doji</td>
<td>270</td>
<td>270</td>
<td>0</td>
</tr>
<tr>
<td>White Paper Umbrella</td>
<td>593</td>
<td>567</td>
<td>0</td>
</tr>
<tr>
<td>Black Paper Umbrella</td>
<td>755</td>
<td>727</td>
<td>1</td>
</tr>
</tbody>
</table>

Panel A: Bullish single lines

Hammer                        | 37    | 57  | 0    | 0.0002 | 0.0003 |
Bullish Engulfing             | 277   | 252 | 0    | 0.0002 | 0.0002 |
Piercing Line                 | 155   | 145 | 0    | 0.0002 | -0.0003 |
Bullish Harami                | 118   | 115 | 1    | 0.0001 | 0.0004 |
Three Inside Up               | 13    | 17  | 0    | 0.0005 | 0.0004 |
Three Outside Up              | 57    | 56  | 0    | 0.0003 | 0.0001 |
Tweezer Bottom               | 348   | 354 | 0    | 0.0002 | 0.0002 |

Panel B: Bullish reversal patterns

Long Black                    | 2670  | 2661 | 0    | -0.0002 | -0.0003 |
Black Marubozu                | 558   | 557 | 0    | -0.0002 | -0.0004 |
Closing Black Marubozu        | 1029  | 1022 | 0    | -0.0002 | -0.0005 |
Opening Black Marubozu        | 1744  | 1737 | 0    | -0.0002 | -0.0002 |
Gravestone Doji              | 192   | 197 | 2    | -0.0002 | -0.0006 |
White Shooting Star           | 544   | 520 | 3    | -0.0004 | 0.0002 |
Black Shooting Star           | 499   | 465 | 1    | -0.0002 | 0.0002 |

Panel C: Bearish single lines

Hanging Man                   | 77    | 84  | 0    | -0.0003 | -0.0005 |
Bearish Engulfing             | 238   | 289 | 0    | -0.0002 | -0.0002 |
Dark Cloud Cover              | 105   | 117 | 1    | -0.0002 | -0.0001 |
Bearish Harami                | 366   | 396 | 0    | -0.0002 | -0.0001 |
Three Inside Down             | 26    | 34  | 0    | 0.0001  | -0.0002 |
Three Outside Down            | 23    | 36  | 0    | -0.0003 | -0.0011 |
Tweezer Top                   | 364   | 407 | 0    | -0.0002 | -0.0003 |

Table 3 contains the full bootstrap results for the GARCH-M model. The number of signals columns contain the number of signals on the bootstrapped (BS) and original series (Dow), respectively. The p-value count columns are the number of stocks for which the mean profit is statistically significantly greater (at the 5% level) on the original series than the 500 bootstrapped series. For a rule to have statistically significant profits on a stock at the 5% level there must be 25 or fewer instances of larger profits to the rules on the random bootstrapped series than the original. Returns are calculated based on the assumption that trades are entered at the open price on the day following a signal, that positions are held for ten days, and that a ten-day exponential moving average is used to determine the prior trend. The Profit column contains the mean return following a candlestick signal on the bootstrapped (BS) and original series (Dow), respectively.

bootstrapped series. For a rule to have statistically significant profits on a stock at the 5% level there must be 25 or fewer instances of larger profits to the rules on the random bootstrapped series than the original series. The Profit column contains the mean profit following a candlestick signal on the bootstrapped (BS) and original series (Dow), respectively.
Bootstrap Buy/Sell are the mean buy/sell profit and across the 500 bootstrapped series, respectively. These are calculated as an average of the 500 series across the 35 stocks. Dow Buy/Sell are the average buy profit across the original series of each of the 35 stocks.

The striking result from the \( p \)-value count columns is the lack of profitability of all the different candlestick rules across the 35 DJIA stocks. Of the 28 candlestick rules, 22 have no statistically significant \( p \)-values on any of the DJIA stocks. The largest number of statistically significant \( p \)-values is 3 for the White Shooting Star. The bullish single lines, bullish reversal patterns and bearish reversal patterns each only have 1 rule that produces statistically significant profits on only 1 of the stocks, while there are only 3 bearish single lines that produce any statistically significant profits.

Columns 5 and 6 contain the mean profit following a candlestick signal on the bootstrap (BS) and original series (orig), respectively. The results in Panel A indicate that profits following all bullish single lines are higher on the random bootstrap series than the original. These results, the opposite to those one would expect, are further evidence of the lack of predictive power of candlestick technical analysis. The relationship between profits on the original and bootstrapped series for bullish reversal patterns is inconsistent. For some rules there are higher mean profits on the original series than the bootstrapped series, while for others the opposite is true. Both bearish single lines and reversal patterns tend to make losses on both the bootstrapped and original series, however, the losses tend to be greater on the original series. This is further evidence against the profitability of candlestick technical analysis.

Candlestick signals are rare compared to typical trading rules considered in the literature (e.g., BLL, 1992) and their forecasting power is only a short-term phenomenon so it is not appropriate to consider daily returns on an annual basis. The daily returns are at times large, but these are not able to be earned over a sustained period of time.

5. Robustness of results

As a further test of the robustness of our results, we investigate the implications of changing some of our key assumptions. This process gives us confidence that our results are not specific to our core assumptions. We find that the sensitivity analysis confirms our early findings. Candlestick technical analysis does not have value for a majority of DJIA stocks under any of the scenarios. The results are very consistent across the scenarios which gives us confidence that they are robust.

Table 4 contains the number of stocks for which the mean profit is statistically significantly greater (at the 5\% level) on the original series than the 500 bootstrapped series. For rules to have statistically significant profits on a stock at the 5\% level there must be 25 or fewer instances of larger profits to the rules on the random bootstrapped series than the original. Once again, the bootstrap results are based on the GARCH-M model. In scenarios A and B we enter a trade at the closing price on the day of (day after) a signal, respectively. Trades are held open for ten days and a ten-day exponential moving average to determine the prior trend for reversal patterns. Scenario C is our base case used in Table 3. This involves entering trades at the open price on the day following a signal, a ten-day holding period, and a ten-day exponential moving average to determine the prior trend. Given that scenario C, with entry at the open price, has very similar results to scenario...
Table 4 contains the bootstrap p-values for GARCH-M model based on different scenarios. The p-value counts are the number of stocks for which the mean profit is statistically significantly greater (at the 5% level) on the original series than the 500 bootstrapped series. For a rules to have statistically significant profits on a stock at the 5% level there must be 25 or fewer instances of larger profits to the rules on the random bootstrapped series than the original. Scenarios A and B are based entering trades at the close price on the day of (day after) a signal, ten-day holding periods, a ten-day exponential average to determine the prior trend. Scenario C is based around entering trades at the open price on the day after a signal, ten-day holding periods, a ten-day exponential moving average. Scenarios D and E are identical to Scenario C but the holding period is changed to five and two days, respectively. Scenarios F and G are identical to Scenario C but the candlestick parameters specifications are changed by +20% and −20%, respectively. Scenarios H and I are identical to Scenario C but five and fifteen day exponential averages, respectively, are used.

B, which is identical except for entry at the close price, the other scenarios consider the impact of changing one of the assumptions in scenario C.

Scenarios D and E are based around holding periods of five and two days, respectively. Scenarios F and G change the candlestick parameter specifications to +20% and −20% of
the original specification, respectively. Scenarios H and I consider five and fifteen day exponential averages, respectively.

The results are very robust to these assumption changes. Across the 28 candlestick rules and 9 assumptions there is weaker profitability on all the 35 DJIA stock series than the bootstrapped series 75% of the time. The maximum number of firms for which candlestick rules are significantly more profitable, than the bootstrapped series, across all candlestick rules and scenario is only three. This is an important result as the assumption changes are at the limits of what can be termed reasonable. While there is some debate on the precise definition of candlestick single lines, changing the core assumptions by +20% and −20% results in specifications that are on the outer limits of those outlined in practitioner books. Because there is no statistical significance, it would be meaningless to consider risk-adjusted returns, transactions costs, and economic significance.

6. Conclusions

The results in this paper indicate that use of the oldest known form of technical analysis, candlestick trading strategies, is not profitable on DJIA stocks over the 1992–2002 period. In contrast to traditional technical analysis, candlestick technical analysis involves analysis of open, high, low and close prices within a day and over successive days.

We propose that the choice of candlestick technical analysis and our choice of data make this study a very robust test of this form of technical analysis. It is less susceptible to the criticisms of data snooping than are many other technical analysis studies. Candlestick technical analysis was developed by Japanese rice traders in the 1700s, therefore testing the technique using DJIA component stock data is an out-of-sample test. The use of a stock data set which is able to be traded in its own right overcomes the criticism that technical analysis profits documented on nontraded indices are purely hypothetical. Individual stock data also overcome any bias introduced by nonsynchronous trading within an index.

By limiting our analysis to the actively traded DJIA stocks we are using prices that could have been obtained by proponents of candlestick technical analysis. The market microstructure of the NYSE means orders could have been filled at the prices we use. Finally, the time frame of our study, 1992–2002, ensures that market participants would have been aware of candlestick technical analysis and had the ability to implement it during this time. This is an important consideration as the challenge to market efficiency from recently developed complex trading rules that are reliant on massive computer power and that are tested on data 50–100 years old is dubious at best.

None of the candlestick rules are found to have forecasting power using an extension of the bootstrapping methodology that accommodates open, high, low and close prices. Trading on the signals generated by candlestick technical analysis therefore does not add value for the major stocks traded in the US market. This evidence is consistent with market efficiency. While it may rational for brokerage firms to include candlestick technical analysis in advice offered to clients if this analysis leads to increased turnover, investors who base their decisions on candlestick technical analysis are unlikely to benefit from it.

Acknowledgements

We thank Jared Cahan and especially Rochester Cahan for outstanding research assistance with the methodology of this paper. Marshall would like to acknowledge the
financial support of the Foundation for Research in Science and Technology (FRST), which was received in the form of a Top Achiever Doctoral Scholarship. We also wish to thank the editor, Professor G.P. Szego, and two anonymous referees for comments that have dramatically improved this paper.

Appendix A. Candlestick single lines and reversal patterns

The description of each candlestick single line and reversal pattern is based around the leading candlestick practitioner books (see footnote 4).

A.1. Bullish single lines

The single lines displayed below are all bullish lines. Each bullish line has a bearish counterpart.

Long White Candle

A Long White Candle, which has a close well above the open towards the high of the day, indicates positive sentiment towards a stock suggesting that the price can be expected to rise in the future.

White Marubozu

A White Marubozu is a long white body with no shadows at either end. This is an extremely strong line as prices have risen throughout the day and closed at their high. It is often the first part of a bullish continuation or bullish reversal candle pattern.

Closing White Marubozu

A Closing Marubozu has no shadow extending from the close end of the body, indicating that prices have closed at their highs. It therefore has similar strength to a Marubozu.

Opening White Marubozu

The Opening Marubozu has no shadow extending from the open price end of the body. The Opening Marubozu is similar to a Long White Candle and not as strong as the Closing Marubozu.

Dragonfly Doji

The Dragonfly Doji occurs when the open and close are at the high of the day. The price declines during the day, but then rallies to close at, or near, the opening price. A Dragonfly Doji at the end of a downtrend is extremely bullish.
White and Black Paper Umbrella

The Paper Umbrella is similar to the Dragonfly Doji. A White Paper Umbrella is strongest as it indicates declining prices throughout the day and then a rally with a close above the opening price. A Black Paper Umbrella is also considered a bullish line as prices have declined throughout the day, but then rallied to close well above their lows. A Black Paper Umbrella is the only black candle that is considered bullish.

A.2. Bearish single lines

The single lines displayed below are all bearish lines.

Long Black Candle

A Long Black Candle, which has a close well below the open towards the low of the day, indicates negative sentiment towards a stock, suggesting that the price can be expected to fall in the future.

Black Marubozu

A Black Marubozu is a long black body with no shadows at either end. This is an extremely weak line as prices have fallen throughout the day and closed at their low. It is often the first part of a bearish continuation or bearish reversal candle pattern.

Closing Black Marubozu

A Closing Marubozu has no shadow extending from the close end of the body, indicating that prices have closed at their lows.

Opening Black Marubozu

The Opening Marubozu has no shadow extending from the open price end of the body. The Opening Marubozu is similar to a Long Black Candle and not as strong as the Closing Marubozu.

Gravestone Doji

The Gravestone Doji occurs when the open and close are at the low of the day. The price rallies during the day, but then declines to close at, or near, the opening price. A Gravestone Doji at the end of an uptrend is extremely bearish.

White and Black Shooting Star

The Shooting Star is similar to the Gravestone Doji. A Black Shooting Star is weakest as it indicates rising prices throughout the day and then a decline with a close below the opening price. A White Shooting Star is also considered a bearish line as prices have risen throughout the day, but then declined to close well below their highs. A White Shooting Star is the only white candle that is considered bearish.
A.3. Bullish reversal patterns

Bullish patterns are defined as those that reduce a bullish single line (i.e., a white candle with a short upper line or a black paper umbrella). Bearish patterns are defined as those that reduce a bearish single line (i.e., a black candle with a short upper line or a white shooting star).

Hammer

The Hammer involves a sell off after a decline to a new intra-day low. Prices then rally to close above the open. Prices on the following day close higher still indicating a reversal has occurred. Nison (1991, p. 29) stated that the lower shadow should be twice the height of the real body and it should have no, or a very short, upper shadow.

Bullish Engulfing

A downtrend must be underway and the first day’s body colour reflects the trend. The second day opens lower, then closes above the open of the first day, indicating a change in sentiment. The Bullish Engulfing pattern reduces to a Hammer which fully supports its interpretation. The Bullish Engulfing pattern is also the first two days of the Three Outside Up pattern.

Piercing Line

The Piercing Line indicates a situation where the market is declining. Following a down day the market opens lower, then rallies throughout the day and closes above the mid-point of the previous day. This action causes concern to bears and indicates that a potential bottom has been made. The Piercing Line is similar to, but not as strong as, the Bullish Engulfing Pattern.

Bullish Harami

Harami is a Japanese word for pregnant or body within. In a Bullish Harami, a long black day perpetuates the downtrend. The next day, prices open higher, which shocks many complacent bears and many short positions are covered causing prices to rise further. This is said to be the first day in a trend reversal.

Three Inside Up

This pattern is a confirmation for the Bullish Harami. Therefore the psychology is the same as that behind the Harami with the added strength that the trend has changed.
The Three Outside Up is confirmation for the Bullish Engulfing Pattern.

Tweezer Bottoms are two or more candlesticks with matching lows. The fact that price is unable to penetrate a given level on successive days indicates that there is good buying support at that level and that the downtrend is likely to reverse.

**A.4. Bearish reversal patterns**

The Hanging Man involves an intra-day decline following an uptrend. Prices then rally, but fail to close above the open. Prices on the following day move lower still, indicating a reversal has occurred. Nison (1991, p. 29) stated that the lower shadow should be twice the height of the real body and it should have no, or a very short, upper shadow.

An uptrend must be underway and the first day’s body colour reflects the trend. The second day opens higher, then closes below the open of the first day, indicating a change in sentiment. The Bearish Engulfing pattern is also the first two days of the Three Outside Down pattern.

The Dark Cloud Cover is a bearish reversal pattern and the counterpart of the Piercing Line pattern. The more penetration of the black body into the prior white body, the greater the chance for a top reversal.

In a Bearish Harami, a long white day perpetuates the uptrend. The next day, prices open lower, which shocks many complacent bulls and many longs are closed causing prices to fall further. This is said to be the first day in a trend reversal.
Three Inside Down

This pattern is a confirmation for the Harami. Therefore the psychology is the same as that behind the Harami with the added strength that the trend has changed.

Three Outside Down

The Three Outside Down is confirmation for the Bearish Engulfing Pattern. The combined pattern reduces to a shooting star which fully supports its interpretation.

Tweezer Top

Tweezer Tops are two or more candlesticks with matching highs. The fact that price is unable to penetrate a given level on successive days indicates that there is good selling resistance at that level and that the down trend is likely to reverse.

References