The predictive power of price patterns

G. CAGINALP and H. LAURENT*

Mathematics Department, University of Pittsburgh, Pittsburgh, PA 15260, USA. E-mail: caginalp@vms.cis.pitt.edu

Using two sets of data, including daily prices (open, close, high and low) of all S&P 500 stocks between 1992 and 1996, we perform a satistical test of the predictive capability of candlestick patterns. Out-of-sample tests indicate statistical significance at the level of 36 standard deviations from the null hypothesis, and indicate a profit of almost 1% during a two-day holding period. An essentially non-parametric test utilizes standard definitions of three-day candlestick patterns and removes conditions on magnitudes. The results provide evidence that traders are influenced by price behaviour. To the best of our knowledge, this is the first scientific test to provide strong evidence in favour of any trading rule or pattern on a large unrestricted scale.

Keywords: candlestick patterns, statistical price prediction, price pattern, technical analysis

1. Introduction

The gulf between academicians and practitioners could hardly be wider on the issue of the utility of technical analysis. On the one hand, technical analysts chart stock prices and carefully categorize the patterns, often with colourful terminology, in order to obtain information about future price movements (see for example Pistolese, 1994). Since the objectives of the technical analysts are highly practical, the rationale and fundamental basis behind the patterns are often subordinated, as is the statistical validation of the predictions. The academicians, on the other hand, are very sceptical of any advantage attained by a method that uses information that is so readily available to anyone. Any successful procedure of this type would violate the efficient market hypothesis (EMH) which implies that changes in fundamental value augmented by statistical noise would be the only factors in a market that would be confirmed by statistical testing. Since there is no inside information involved in charting prices, a statistically valid method would violate all forms of EMH. Any scheme that is valid, many economists would argue, would soon be used by many traders, the advantages would diminish and the method would self-destruct. In the absence of clear, convincing statistical evidence in a highly scientific and objective form, it is not surprising that technical analysis is dismissed so readily (e.g. Malkiel, 1995). Brock *et al*. (1992) review the literature of the past few decades and also conclude that most studies have found no statistical validity in trading rules that were tested. Brock *et al*. study two trading rules, including a moving average test which is defined by a 'buy signal' when prices cross a moving average (of 200 days, for example) on the upside and a 'sell signal' when prices cross on the downside. Their highly

*Current addresss: Grant Street Advisors, 429 Forbes Avenue, Pittsburgh, PA 15219, USA

sophisticated statistical methods detected only a slight positive net gain upon utilization of the moving average methods when tested against the Dow Jones Industrial Average. In fact a statistically significant loss appeared for the sell signal that may be attributed to the predictability involved in the volatility.

The additional problem that technical analysis encounters relates to basic microeconomic issues. While the fundamental justification for charting tends to be skimpy, the objections summarized above would be addressed by technical analysts as follows. The large influx of investors and traders in the world's markets make it unlikely that the vast majority would be able to utilize a complicated set of methods or find skilful money managers who are able to do it for them. Another factor that complicates the EMH argument is the frequent lack of agreement among experts on the true or fundamental value of an asset. Publications such as the *Wall Street Journal* and *Barron's* frequently list the predictions of the major players on currencies, for example, and the variation in the predictions for a year hence often exceed the variation in actual prices during the past year. It is not difficult to find two leading investment houses, of which one feels a currency or index is overvalued while the other feels it is undervalued. Thus, the temporal evolution of prices will reflect not only fundamentals, but the expectations regarding the behaviour and assets of others. Also, many investors would find it difficult to ignore the changes in price of their asset, and refrain from selling in a declining market. An academic study that has provided a mathematical explanation of technical analysis from this perspective is Caginalp and Balenovich (1996) which demonstrated that technical analysis patterns could be obtained as a result of some basic assumptions involving trend based investing, finite resources and in some cases asymmetric information. Another microeconomic derivation by Blume *et al*. (1994) has suggested that traders might obtain additional insight into the direction of prices by utilizing the information in the trading volume.

To any practitioner, it is evident that academics routinely overestimate the amount of money available to capitalize quickly on market inefficiencies and the availability of honest, reliable, and inexpensive advice on a market. Many fund managers, for example, are constrained to low-turnover in their funds and are highly restricted in trading for their own accounts. There are also implicit political constraints against advising the sale or short-sale of securities. Thus the pool of 'smart' money available to speculate is not as large as it might seem. A related issue is whether the professionals really compete with one another and thereby render the market efficient, or effectively exploit the less sophisticated investors by declining to trade unless there is a rather predictable, healthy profit to be made, thereby leaving the market in a less efficient state. Without any assurance that the price will necessarily gravitate toward a realistic value (which is not unique as discussed above) the trading decisions will thereby focus upon the evolution of the strategies of others as they are manifested in the price behaviour of the asset. A great deal of insight into this issue has been provided through economics experiments (Porter and Smith, 1994) that found bubbles persisted under very robust conditions, despite the absence of any uncertainty in the asset traded. One of the few changes to experimental design that eliminated the bubbles was the attainment of experience as an integral group of traders.

A discussion of EMH at this level provokes a more fundamental question on the motivation of a trader who perceives a discount from fundamental value. The purchase of the asset ordinarily does not guarantee a profit except as a consequence of optimization by others. The issue of distinguishing between self-maximizing behaviour and reliance on the optimizing behaviour of others is considered in the experiments of Beard and Beil (1994) on the Rosenthal conjecture (1981) that showed the unwillingness of agents to rely on others' optimizing behaviour. In these experiments, player A can

choose a smaller payout that does not depend on player B, or the possibility of a higher payout that is contingent on player B making a choice that optimizes B's return (otherwise A gets no payoff). The experiments showed that A will accept the certain but smaller outcome that is independent of B. However, reinforcing a line of reasoning that is compatible with Meyerson's (1978) 'proper equilibrium' in which 'mistakes' are related to the payoff consequences, Beard and Beil (1994) showed that player A will be less reluctant to depend on B when 'deviations from maximality become more costly for B' (p. 257). Thus, the strategies of other players clearly provide important information relating future prices, and recent market history is the only inkling one has on these strategies. Thus, one may conclude that price action (and other market information) provides important information unless, as EMH advocates may assert, the extent of use among traders renders them useless. This question can only be answered by a large-scale statistical test, which is the subject of this paper.

In view of the discussion above, the statistical validity or repudiation of basic charting techniques takes on a significance beyond the issue of immediate profitability. If a fair test demonstrates statistical significance of basic charts then it would certainly refute the key claim of EMH that any successful method immediately sows the seeds of its own destruction. Furthermore, it would add to the evidence provided by the experiments that traders are keenly focused on the actions of others as they are manifested through price action in making their investment decisions. Finally, it would mean that markets contain more than random fluctuations about fundamental value, so that it is meaningful to investigate the remaining deterministic forces. Analogously, the lack of statistical significance in a fair test would provide further evidence of market efficiency and the underlying assumptions inherent in it.

The extent of efficiency in markets has been studied within the context of statistical models in a number of studies which have obtained mixed conclusions about this key question (Lo and MacKinley, 1988; Shiller, 1981; White, 1993).

A standard method of testing for market efficiency is to embed it as a linear autoregressive model for asset return, r_t, at time t, of the form

$$r_t = a_0 + a_1 r_{t-1} + \ldots + a_{t-p} r_{t-p} + \varepsilon_t$$

for some $p \in Z^+$ where $(a_0, \ldots, a_p)$ is an unknown vector of coefficients. As noted in White (1993) evidence of the form $a_1 \neq 0$, $a_2 \neq 0$, $\ldots$, $a_{t-p} \neq 0$ is contrary to the assertions of EMH. However, empirical evidence that $a_1 = a_2 = \ldots = a_{t-p} = 0$ does not establish EMH entirely since it has been shown that one can have deterministic nonlinear processes that possess no linear structure (Brock, 1986).

In the study by White (1993) on the possibility of using neural networks to predict the price of IBM stock, it is demonstrated that the coefficients do not deviate significantly from zero. However, an ARIMA study by Caginalp and Constantine (1995) on a quotient of two 'clone' closed-end funds (Germany and Future Germany) found a large coefficient indicating a strong role for price momentum once exogenous random events (that influence the overall German Market) are removed in this way. The White (1993) study makes a pessimistic conclusion about very simple models of neural networks for finance. Part of the problem is that one may require a long 'training period'. Another is that deeper levels of networks may result in 'overfitting'. Using conventional technical analysis patterns may avoid these problems in that the role of experience of traders provides a long history on which the training period occurs.

A fair test of charting, however, encounters several problems.

(1) The definition of the pattern is often not percise in a scientific sense.
(2) Some patterns take weeks to develop, so that random events influencing fundamentals may make testing difficult.

In developing a fair test of technical analysis we focus on some short-term indicators that avoid these difficulties. It is worth noting that there has been some progress on addressing (1) in the case of 'triangle' patterns by Kamijo and Tanigawa (1993). As described in Section 2, we consider a technique known as Japanese candlesticks that has the following advantages.

(a) The definitions tend to be more precise than in the longer patterns.
(b) The time intervals are fixed, facilitating statistical tests.
(c) The method has been in use for many years so that it confronts directly the issue of whether a simple method will self-destruct in a short time due to overuse.

We perform statistical tests to determine whether the appearance of a set of patterns changes the probability that prices will rise or fall and, moreover, whether trading based on these patterns will be profitable. Due to the characteristics listed above it is possible to do this almost completely non-parametrically. Morris (1992) tabulates some consequences of implementing a strategy based on a large number of candlestick patterns with mixed results. Since the key issues of the definition of a trend and its demise are based on visual observations, the study is difficult to evaluate on strict scientific criteria. A key difference between our work and Morris (1992) is that we isolate a specific timescale for the pattern and its effect. The basic ideas of candlestick patterns are discussed in Section 2 while the statistical test is described in Section 3. The conclusions are summarized in Section 4.

2. Candlestick patterns

2.1 *The history of Japanese candlestick patterns*

A form of technical analysis known as Japanese candlestick charting dates back to 18th century Japan when a man named Munehisa Honma attained control of a large family rice business. His trading methodology consisted of monitoring the fundamental value (by using over 100 men on rooftops every four kilometres to monitor rice supplies) as well as the changing balance of supply and demand on the marketplace by tracking the daily price movements. The technical aspect of the analysis is based on the premise that one can obtain considerable insight into the strategies and predicaments of other players by understanding the evolution of open, close, high and low prices. While Western analyses have traditionally emphasized the daily closing prices as the most significant in terms of commitment to the asset, the use of candlesticks has been increasingly popular since Steve Nison introduced it to American investors in the 1970s (Nison, 1991).

Candlestick analysis has been developed into a more visual and descriptive study over the years (Fig. 1). Each candlestick (black or white) represents one trading day (Fig. 2). The white candlestick opens at the bottom of the candle and closes at the top, while the black candlestick is the reverse. In both cases the lines above and below represent the trading range. If the close and open are equal for a particular day, then the 'body' of the candlestick collapses into a single horizontal line.
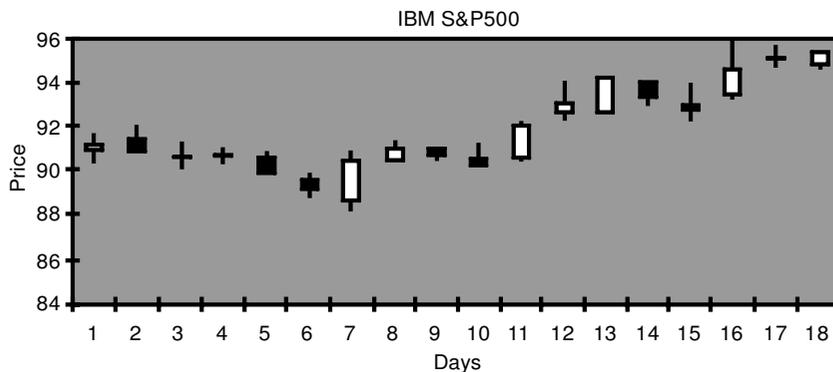
Fig. 1. Sample candlestick chart.

Fig. 2. Candlesticks representing one trading day

The Japanese candlestick method comprises many patterns with differing time scales (usually between one and three days) that offer various levels of confidence. Morris (1992), for example, discusses each of the patterns in terms of whether 'confirmation' is necessary. In keeping with Honma's philosophy, most of the reliable patterns ('no confirmation necessary') are expected to be the patterns that occur over a three-day (or longer) period. The rationale for this is that a manifestation of trading strategies that occurs within a shorter period may not reflect accurately the changing balance of supply and demand, but rather a momentary change in sentiment.

Our study focuses on eight (non-overlapping) three-day 'reversal' patterns that are tested in terms of their ability to forecast a change in the direction of the trend. In making the definitions, we simplify the interpretive aspects of the traditional definitions. For example, the condition of being a 'long' day, i.e. that the magnitude of the open minus the close is large, will be omitted. We do this in order to maintain nonparametric testing, with the expectation that omission of this condition would reduce but not eliminate a statistically positive result, if in fact there is substance to this methodology. Furthermore, the patterns often refer to 'uptrend' or 'downtrend'. These concepts are crucial to the idea of candlesticks. In other words, a three-day pattern itself without the correct trend is irrelevant as an indicator. Consequently, we must make a suitable definition of downtrend. We do this by smoothing out the daily closing prices using a three-day moving average and then requiring that the

moving average is decreasing in each of the past six days except possibly one (see Definition 3.1). In all of our mathematical definitions of candlestick patterns, the term uptrend or downtrend will utilize this definition of trend. Of course, in traditional technical analysis, the term trend is used in a more vague sense based upon visual observation, though any two technical analysts looking at the same graph with the same timescale are likely to agree on where the trends appear. Once again, our nonparametric definitions would pick up very small trends that would be negligible in practice, thereby diluting the statistical results. However, the alternative would be the use of parameters to define the magnitude of the trend, which we are seeking to avoid. As a consequence of our definitions, then, the statistical tests we define will have a slight built-in bias against the methodology. Our approach is completely out-of-sample, since the definitions are formulated largely in Morris (1992) which uses data sets that are from a previous time period, and usually for commodities futures.

2.2 Pattern definitions

We label each of the three consecutive days of which the test will take place as $t^* + 1$, $t^* + 2$, and $t^* + 3$, as shown in Fig. 3.
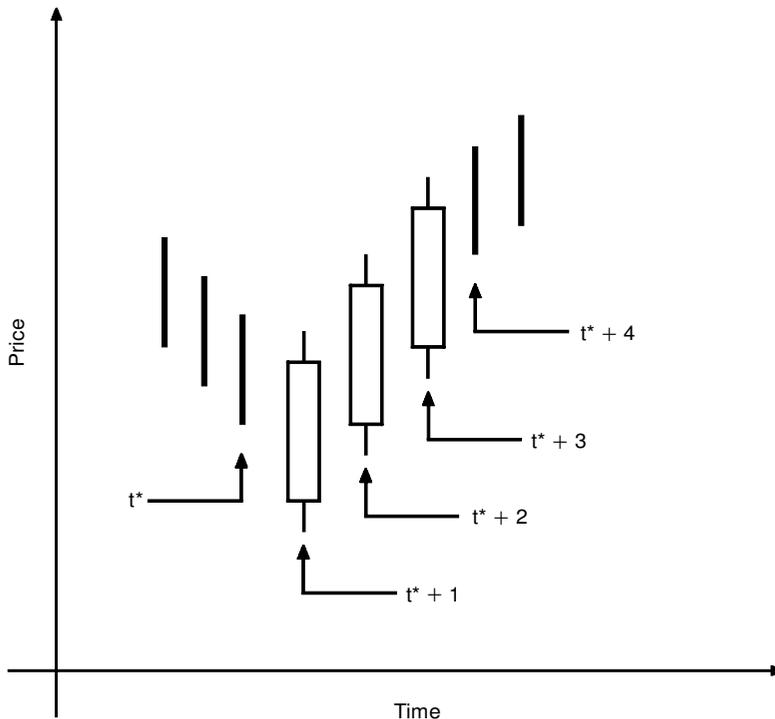

Fig. 3. Three day candlestick patterns

2.2.1 *Three White Soldiers (TWS)*

As explained in Morris (1992), the Three White Soldiers pattern is composed of a series of long white candlesticks which close at progressively higher prices and begin during a downtrend. The hypothesis is that the appearance of a Three White Soldiers is an indication that the downtrend has reversed into an uptrend (Fig. 4). In order to make this definition mathematically precise and nonparametric, we reformulate it using Definition 3.1 for the downtrend and eliminate the condition on the length of the candlestick body.

Definition (Three White Soldiers)
(1) The first day of the pattern, $t^* + 1$, belongs to a downtrend in the sense of Definition 3.1.
(2) Three consecutive white days occur, each with a higher closing price:

$$c_i - o_i > 0 \text{ for } i = t^* + 1, \, t^* + 2, \, t^* + 3$$

$$c_{t^*+3} > c_{t^*+2} > c_{t^*+1}$$

(3) Each day opens within the previous day's range

$$c_{t^*+1} > o_{t^*+2} > o_{t^*+1}$$

$$c_{t^*+2} > o_{t^*+3} > o_{t^*+2}$$

Fig. 4. Three white soldiers.

2.2.2 *Three Black Crows (TBC)*

The Three Black Crows pattern is the mirror image of the three white soldiers. It usually occurs when the market either approaches a top or has been at a high level for some time, and is composed of three long black days which stair-steps downward. Each day opens slightly higher than the previous days close, but then drops to a new closing low. TBC is a clear message of a trend reversal (Fig. 5).

Definition (Three Black Crows)

(1) The first day of the pattern, $t^* + 1$, belongs to an uptrend in the sense of Definition 3.1.

(2) Three consecutive black days occur, each with a lower closing price:

$$o_i - c_i > 0 \text{ for } i = t^* + 1, \, t^* + 2, \, t^* + 3$$

$$o_{t^*+1} > o_{t^*+2} > o_{t^*+3}$$

$$c_{t^*+1} > c_{t^*+2} > c_{t^*+3}$$

(3) Each day opens within the previous day range:

$$o_{t^*+1} > o_{t^*+2} > c_{t^*+1}$$

$$o_{t^*+2} > o_{t^*+3} > c_{t^*+2}$$
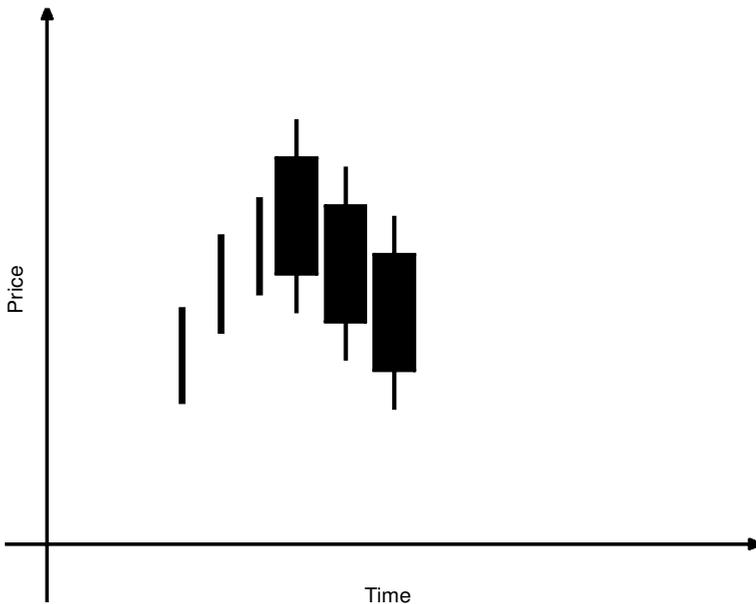

Fig. 5. Three black crows.

2.2.3 *Three Inside Up (TIU)*

A Three Inside Up pattern (Morris, 1992) occurs when a downtrend is followed by a black day that contains a small white day that succeeds it. The third day is a white candle that closes with a new high for the three days (Fig. 6).

Definition (Three Inside Up)
(1) The first day of the pattern, $t^* + 1$, belongs to a downtrend in the sense of Definition 3.1.
(2) The first day of pattern, $t^* + 1$, should be a black day:

$$o_{t^*+1} > c_{t^*+1}$$

(3) The middle day, $t^* + 2$, must be contained within the body of the first day of the pattern $t^* + 1$:

$$o_{t^*+1} \geqslant o_{t^*+2} > c_{t^*+1}$$

$$o_{t^*+1} > c_{t^*+2} \geqslant c_{t^*+1}$$

with at most one of the two equalities holding. That is, the opening prices or the closing prices of the two days may be equal but not both. Hence, either the open or close (but not both) of the $t^* + 1$ and $t^* + 2$ days may be equal.
(4) Day $t^* + 3$ has a higher close than open and closes above the open of day $t^* + 1$.
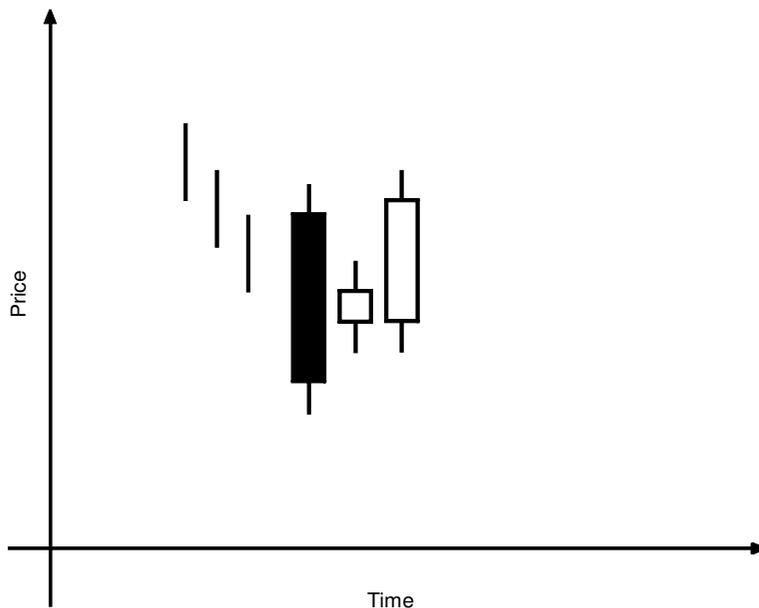
$$c_{t^*+3} > o_{t^*+3}$$

$$c_{t^*+3} > o_{t^*+1}$$

Fig. 6. Three inside up.

2.2.4 *Three Inside Down (TID)*

The Three Inside Down pattern is the topping indicator analogous to the three inside up pattern (Fig. 7).

Definition (Three Inside Down)
(1) The first day of the pattern, $t^* + 1$, belongs to an uptrend in the sense of Definition 3.1. The first day, $t^* + 1$, has a higher close than open.

$$c_{t^*+1} - o_{t^*+1} > 0$$

(2) The middle day, $t^* + 2$, must be contained within the body of the first day of the pattern $t^* + 1$:

$$c_{t^*+1} > o_{t^*+2} \geqslant o_{t^*+1}$$

$$c_{t^*+1} \geqslant c_{t^*+2} > o_{t^*+1}$$

which at most one of the two equalities holding. That is, the opening prices or the closing prices of the two days may be equal but not both. Hence, either the open or close (but not both) of the $t^* + 1$ and $t^* + 2$ days may be equal.
(3) The third day, $t^* + 3$, has a lower close than open, and its close is lower than the first day's open:

$$o_{t^*+3} - c_{t^*+3} > 0$$
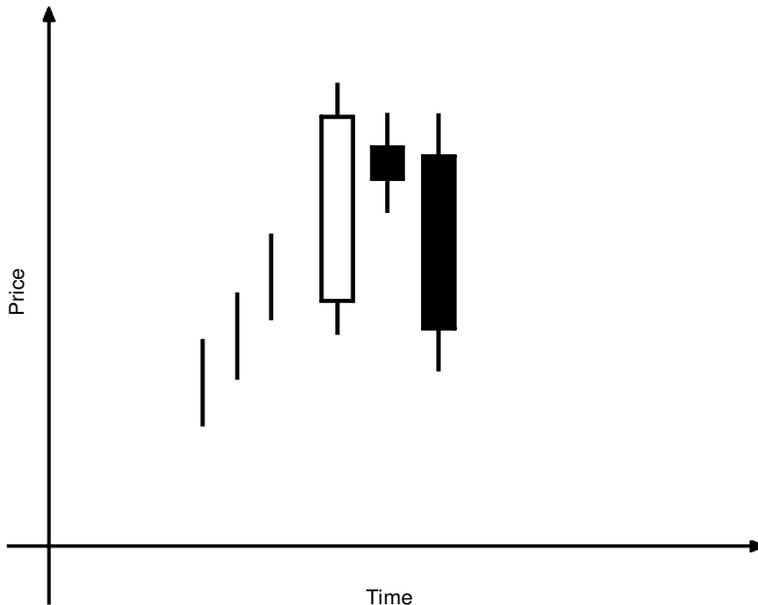
$$c_{t^*+3} > o_{t^*+1}$$

Fig. 7. Three inside down.

2.2.5 *Three Outside Up (TOU)*

The Three Outside Up is similar to the Three Inside Up, with the second day's body engulfing the first day's body amid rising prices. The third day, a white candle, closes with a new high for the three days, giving support to this reversal (Fig. 8).

Definition (Three Outside Up)
(1) The first day of the pattern, $t^* + 1$, belongs to a downtrend in the sense of Definition 3.1 and has a higher open than close:

$$o_{t^*+1} - c_{t^*+1} > 0$$

(2) The second day $t^* + 2$ must completely engulf the prior day, $t^* + 1$ in the sense of the following inequalities:

$$c_{t^*+2} \geqslant o_{t^*+1} > c_{t^*+1} \geqslant o_{t^*+2}$$

$$|c_{t^*+2} - o_{t^*+2}| > |c_{t^*+1} - o_{t^*+1}|$$

(3) The third day, $t^* + 3$, has a higher close than open, and closes higher than the second day, $t^* + 2$:

$$c_{t^*+3} - o_{t^*+3} > 0$$
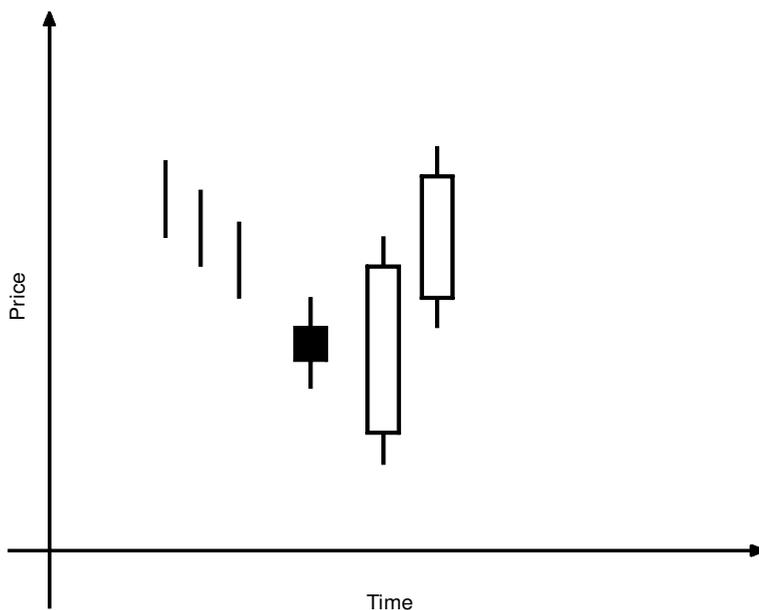
$$c_{t^*+3} > c_{t^*+2}$$

Fig. 8. Three outside up.

2.2.6 *Three Outside Down (TOD)*

The Three Outside Down pattern is the up-to-down reversal pattern analogous to TOU (Fig. 9).

Definition (Three Outside Down)
(1) The first day of the pattern, $t^* + 1$, belongs to an uptrend in the sense of Definition 3.1. The first day also has a higher close than open:

$$c_{t^*+1} - o_{t^*+1} > 0$$

(2) The second day $t^* + 2$, a black day, must completely engulf the prior day $t^* + 1$ in the sense of the following inequalities:

$$o_{t^*+2} \geqslant c_{t^*+1} > o_{t^*+1} \geqslant c_{t^*+2}$$

$$|c_{t^*+2} - o_{t^*+2}| > |c_{t^*+1} - o_{t^*+1}|$$

(3) The third day $t^* + 3$ is a black candle with a lower close than the previous day:

$$o_{t^*+3} - c_{t^*+3} > 0$$
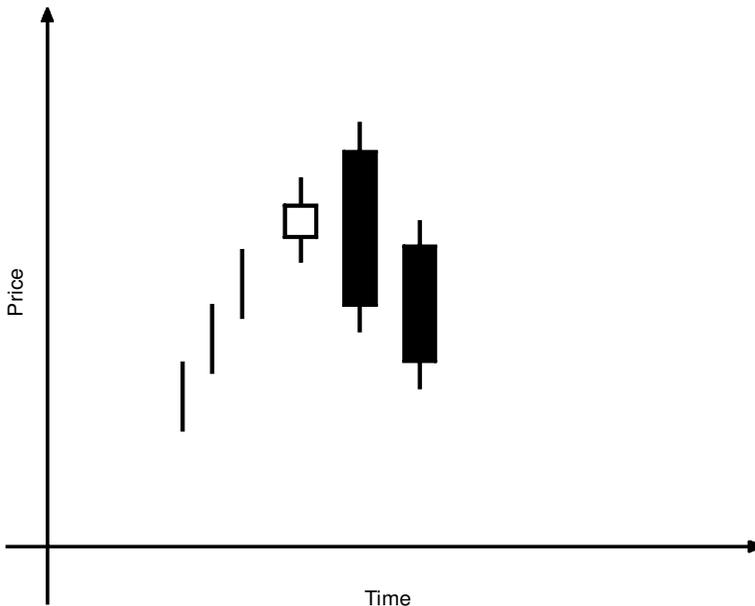
$$c_{t^*+3} < c_{t^*+2}$$

Fig. 9. Three outside down.

2.2.7 *Morning Star (MS)*

This pattern forms as a downtrend continues with a long black day. The downtrend receives further confirmation after a downward gap occurs the next day. However, the small body, black or white, shows the beginning of market indecision (or some indication that supply and demand have become more balanced). Prices rise during the third day, closing past the midpoint of the first day's body (Fig. 10), signalling a reversal.

Definition (Morning Star)
(1) The first day, $t^* + 1$, is black and belongs to a downtrend market in the sense of Definition 3.1:

$$o_{t^*+1} - c_{t^*+1} > 0$$

(2) The second day, $t^* + 2$, must be gapped from the first day, and can be of either colour:

$$|o_{t^*+2} - c_{t^*+2}| > 0$$

$$c_{t^*+1} > c_{t^*+2} \text{ and } c_{t^*+1} > o_{t^*+2}$$

(3) The third day $t^* + 3$, is a white day, and ends higher than the midpoint of the first day, $t^* + 1$:

$$c_{t^*+3} - o_{t^*+3} > 0$$

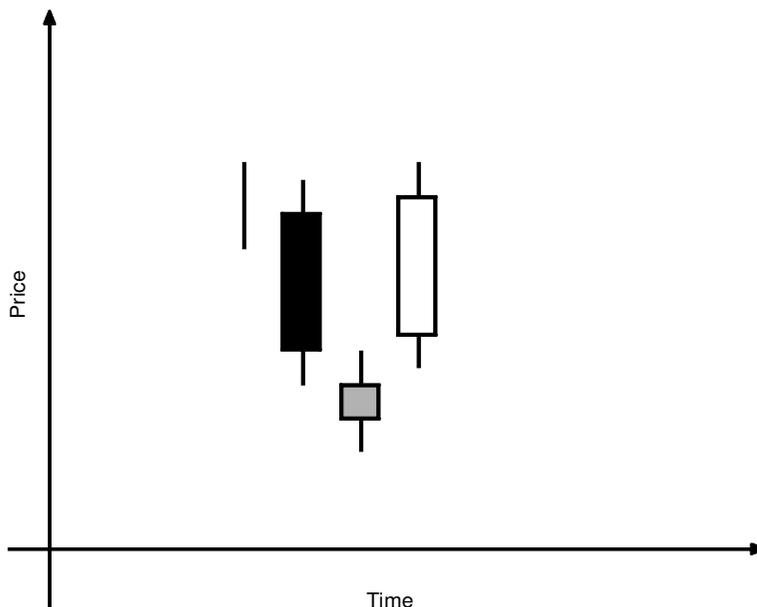$$c_{t^*+3} > \frac{o_{t^*+1} - c_{t^*+1}}{2}$$

Fig. 10. Morning star.

2.2.8 *Evening Star (ES)*

The Evening Star is the mirror image of the Morning Star. It signals a reversal from an uptrend to a downtrend (Fig. 11).

Definition (Evening Star)
(1) The first day, $t^* + 1$, of the pattern belongs to an uptrend and is white day:

$$c_{t*+1} - o_{t*+1} > 0$$

(2) The second day $t^* + 2$ is gapped from the first day body amd can be of either colour. However the open and close of the second day cannot be equal:

$$|o_{t*+2} - c_{t*+2}| > 0$$

$$c_{t*+2} > c_{t*+1} \text{ and } o_{t*+2} > c_{t*+1}$$

(3) The third day $t^* + 3$, is black and ends lower than the midpoint of the first day ($t^* + 1$):

$$o_{t*+3} - c_{t*+3} > 0$$
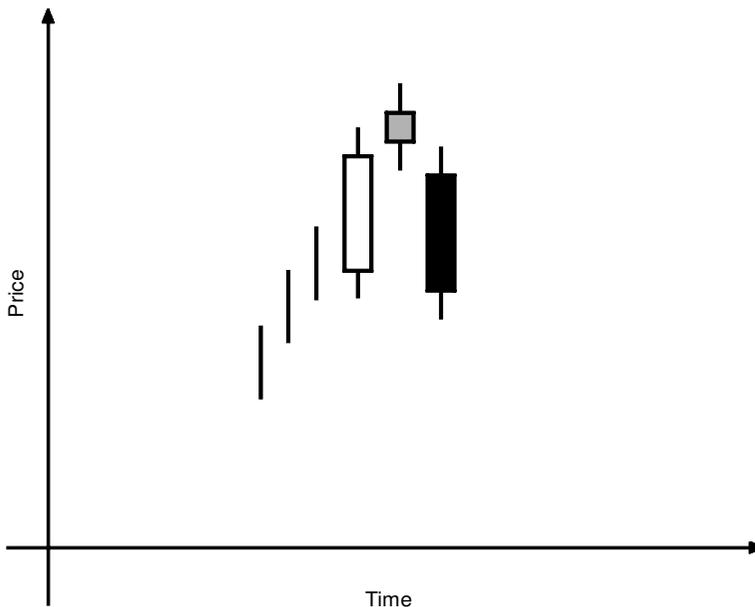
$$c_{t*+3} < \frac{c_{t*+1} - o_{t*+1}}{2}$$

Fig. 11. Evening star.

3. Test of hypothesis

The central objective is to determine whether the candlestick reversal patterns have any predictive value. The reversal patterns are expected to be valid only when prices are in the appropriate trend. Formulating a suitable mathematical definition of trend is a delicate issue, since those given by technical analysts often make use of 'channels' that would be highly parametric in nature and subject to interpretation. Consequently, we make a definition that is essentially nonparametric except for the time scale, with the expectation that the essence of the concept will be captured with only a slight bias against the validity of candlesticks.

The three-day moving average at time t is defined by:

$$M_{avg}(t) = \frac{1}{3}\{P(t-2) + P(t-1) + P(t)\}$$

where $P(t)$ denotes the closing price on day t.

Definition 3.1
A point t is said to be in a *downtrend* if

$$M_{avg}(t-6) > M_{avg}(t-5) > \ldots > M_{avg}(t)$$

with at most one violation of the inequalities. *Uptrend* is defined analogously.

This captures the general idea that the prices are tending downward but allows for the possibility of fluctuation. The time period of six days corresponds to two lengths of the basic patterns. While there is some arbitrariness in this definition, it is one of the two instances where a parameter has been used, and robustness will be checked in both cases.

For concreteness, we focus on downtrends as the issues are identical for uptrends. For a particular stock, suppose that t^* is in a downtrend in the sense of Definition 3.1. The hypothesis we would like to test is that the existence of a candlestick reversal pattern such as TWS increases the likelihood of prices moving higher. To be more precise we use $P(t^*)$ to denote the closing price on day t^* and determine whether the following statement $(A1)$ is true.

$$P(t^* + 3) \leqslant P_{avg}(t^* + 4, t^* + 5, t^* + 6) \tag{A1}$$

Here $P_{avg}(t^* + 4, t^* + 5, t^* + 6)$ is simply the average of the closing prices on those days. Note that we avoid using $P_{avg}(t^* + 3)$ on the left-hand side since this would simply confirm what we know, e.g. for TWS, that prices had been lower. The use of $P_{avg}(t^* + 4, t^* + 5, t^* + 6)$ instead of simply $P(t^* + 4)$ provides a smoothing of the data within the time scale under consideration.

To check for robustness of the results we also vary $(A1)$ within the same general time scale to formulate conditions $(A2)$, $(A3)$ and $(A4)$ below

$$P(t^* + 3) \leqslant P_{avg}(t^* + 4, t^* + 5) \tag{A2}$$

$$P(t^* + 3) \leqslant P_{avg}(t^* + 5, t^* + 6) \tag{A3}$$

$$P(t^* + 3) \leqslant P_{avg}(t^* + 5, t^* + 6, t^* + 7) \tag{A4}$$

To test for predictive power, the first step is to establish the overall probability (with or without the

Table 1. Statistics for the World Equity Closed-Ends Funds.

	Equation	A1	A2	A3	A4	B1	B2	B3	B4
Overall number	$\dfrac{n_0 p_0}{n_0}$	11803/ 26386	14149/ 26386	14534/ 26386	11985/ 26386	13672/ 24380	14001/ 24380	14260/ 24380	13444/ 24380
Percent	p_0	44.73%	53.62%	55.08%	45.42%	56.07%	57.42%	58.49%	55.14%
With Candlestick reversal	$\dfrac{np}{n}$	88/140	94/140	91/140	90/140	188/280	181/280	187/280	190/280
Percentage	p	62.85%	67.14%	65.00%	64.28%	67.14%	64.28%	66.78%	67.85%
Expected number	np_0	62	75	77	63	156	160	163	154
Standard deviation	$\sigma = \sqrt{np_0(1-p_0)}$	5.85	5.89	5.88	5.86	8.27	8.25	8.22	8.31
Number of standard deviations away from the null hypothesis	$Z = \dfrac{n(p-p_0)}{\sigma}$	4.33	3.21	2.36	4.5	3.74	2.32	2.82	4.28

Of the n_0 points t^* are in a downtrend, a fraction p_0 satisfy A1, namely $P(t^*+3) < P_{avg}(t^*+4, t^*+5, t^*+6)$. If in addition there is a candlestick reversal pattern in points (t^*+1, t^*+2, t^*+3) then there are total of n points of which a fraction p satisfy A1, deviating from the null hypothesis $(p = p_0)$ by 4.33 standard deviations. The analogous situation for an uptrend is described by B1 and 3.74 standard deviations is obtained. The conditions A2, A3, A4, and B2, B3, B4 are modifications of A1 and B1 respectively that establish robustness.

candlestick patterns), p_0, for which statement ($A1$) is valid among those t^* that are in a downtrend. In Data Set 1 (Table 1), which consists of daily prices (open, close, high and low) of all world equity closed end funds (as listed in *Barron's*) during the period 4/1/92 to 6/7/96 that were available with sufficient data (54 in all). In all 26386 points were found to be in a downtrend (all stocks combined) and 11803 of those satisfy condition ($A1$) so that $p_0 = 11803/26386 = 44.73\%$. This establishes the mean, which due to the large sample of 26386, has sufficiently small standard deviation that we can assume it is the hypothetical mean. (This reduces the statistical analysis to examining the mean of a single sample population). The next step is to determine the number, n, of points, t^* satisfying not only the condition of being in a downtrend but also the condition that ($t^* + 1$, $t^* + 2$, $t^* + 3$) are a (down-to-up) candlestick reversal pattern, e.g. TWS. Within this subpopulation we determine the fraction, p, and the number, np, for which ($A1$) is true. For Data Set 1, one finds $n = 140$ and $p = 62.85\%$ and $np = 88$. The statistical significance of the deviation of the mean can be computed using the central limit theorem so that a normal distribution can be assumed and the standard deviation is given by

$$\sigma = \sqrt{np_0(1 - p_0)}$$

We note that the points are not completely independent with respect to satisfying Definition 3.1 (or any reasonable definition of a trend), however, the correlations are very small since the moving average involves relatively few points compared to the sample sizes, so the estimate of the standard deviations is reasonably accurate.

The difference between the two means np_0 (the expected number of 'successes') and np (the actual number) is measured in number of standard deviations from the null hypothesis by

$$Z = \frac{n(p - p_0)}{\sigma}$$

For Data Set 1 (Table 1), one obtains the expected number as $np_0 = 62$ and $Z = 4.433$. This provides very strong evidence that these reversal patterns provide a statistically significant indication of a change in the trend.

This procedure is repeated using each of the conditions ($A2$), ($A3$) and ($A4$) in place of ($A1$). The results, shown in Table 1, are similar with $Z = 3.21$, $Z = 2.36$, $Z = 4.5$, respectively. Similarly, one can vary the definition of trend without much change in the Z values, so that robustness is confirmed.

The analogous results for up-to-down reversals are summarized in Table 1. In particular, for ($B1$) the overall mean is $p_0 = 56.07\%$ while the candlestick property mean $p = 67.14\%$ with $\sigma = \sqrt{156(0.5607)(0.4393)} = 8.27$, $np_0 = 156$ and $Z = 3.74$. The results for ($B2$), ($B3$) and ($B4$) are very similar.

Similar tests are performed on Data Set 2, which consists of daily prices (open, close, high and low), of all stocks in the 1996 listing of the Standard and Poor's 500 during the time period 2 January 1992 to 14 June 1996, for which data were available through the commercial service in use. (Note that a handful of stocks of the S&P 500 in 1996 had too short a price history, due to mergers or other changes, and were omitted from the study at the outset.) Performing the same tests on the much larger set of data, we obtained slightly better percentages and an astronomical set of Z values, which yield a high statistical significance even when one compensates for dependencies.

In particular, ($A1$) is valid with probability $p_0 = 45.05\%$ in the overall data but with $p = 71.22\%$

for the set satisfying the candlestick pattern criteria. This implies a statistical significance at the level of $Z = 36.03$ standard deviations away from the null hypothesis.

The results for uptrend reversals indicate that $(B1)$ is satisfied with probability $p_0 = 52.78\%$ within the set of points that is in an uptrend and with $p = 67.33\%$ within the set of points that is in an uptrend and also satisfy the candlestick reversal pattern criteria. The statistical significance is at the level of 26.7% standard deviations. The checks for robustness, displayed under $(A2)$, $(A3)$, $(A4)$ and $(B2)$, $(B3)$, $(B4)$ in Table 2 show similar deviations from the null hypothesis. The percentage of successful reversals [(Ai) or (Bi) is true] appears to deviate by relatively small amounts, namely 71.22% to 73.69% for the (Ai) and 66.75% to 57.50% for the (Bi). Consequently, the predictive power appears to be very robust. A modification of the definition of moving average, e.g. from a three-day to a four-day moving average, makes little difference, as does a similar change in the definition of the trend. These robustness checks thereby reduce the influence of the few parameters that have been used in the system. Of course, the concept of a time scale is intrinsic to this type of short-term indicator, so that the predictive power can be expected to disappear as one takes very large time scales for the trend, the moving average and the predicted closing price (such as $(A2)$, $(A3)$, etc.).

Remark 3.2

The eight candlestick patterns include as their first condition that the first day of the pattern, $t^* + 1$ be part of a downtrend in the sense of Definition 3.1. In calculating the overall probabilities, p_0, of points satisfying (Ai) we require that t^* be part of a downtrend. Consequently, we are requiring a bit more in terms of the trend for the candlestick patterns which must be regarded as part of the definition. Since there is a great deal of averaging in the definition of the trend, the statistical difference arising from requiring t^* instead of $t^* + 1$ to be in a downtrend is likely to be very small. Nevertheless, we can examine this by comparing the p_0 obtained from $(A1)$, which examines the effect of t^* belonging to a downtrend on the points $t^* + 4$, $t^* + 5$, $t^* + 6$, with p obtained from $(A4)$, in which the effect of the $t^* + 1$ point on the points $t^* + 5$, $t^* + 6$, $t^* + 7$ is measured. This makes a very slight change, as the original comparison of (p_0, p) in $(A4)$ is changed from (45.72%, 72.58%) to (45.05%, 72.05%), making the result even stronger in the case of the S&P 500 data. In this comparison, the last point in the pattern is one more day removed from the pattern, demonstrating the robustness of the result. Similarly, the p_0 obtained from $(A2)$ can be compared with p in $(A3)$, so that the comparison of (p_0, p) is changed from (53.57%, 73.40%) to (52.60%, 73.40%), again making the result slightly stronger. Similar results are obtained in comparing the uptrend results (Bi), and for the World Equity Closed-End Fund data.

Remark 3.3

The S&P 500 contains groups of stocks that are correlated, e.g. the banking sector, the computer sector, etc., so that there is some dependence among the daily price movements of the 500 stocks. Consequently, the Z values obtained are overstated to some extent. Given the large differences obtained for p_0 and p in each case, we can easily obtain a lower bound for the statistical significance by underestimating the number of 'independent' stocks. If we assume that there are 50 groups that are completely correlated this reduces the number of stocks by a factor of 10. Furthermore, if we reduce the number of data points, n, by an additional factor of 3 to compensate for dependence due to possible time overlap in the patterns, we reduce the value of n by a factor of

Table 2. Statistics for the S&P 500 stocks.

	Equation	A1	A2	A3	A4	B1	B2	B3	B4
Overall number	$\dfrac{n_0 p_0}{n_0}$	119678/ 265648	139746/ 265648	142321/ 265648	121472/ 265648	132461/ 250923	140101/ 250923	142759/ 250923	134470/ 250923
Percent	p_0	45.05%	52.60%	53.57%	45.72%	52.78%	55.83%	56.89%	53.50%
With Candlestick reversal	$\dfrac{np}{n}$	3339/ 4688	3455/ 4688	3441/ 4688	3403/ 4688	5650/ 8391	5601/ 8391	5622/ 8391	5664/ 8391
Percentage	p	71.22%	73.69%	73.40%	72.58%	67.33%	66.75%	67.00%	67.50%
Expected number	np_0	2111	2465	2511	2143	4428	4684	4773	4489
Standard deviation	$\sigma = \sqrt{np_0(1 - p_0)}$	34.05	34.18	34.14	34.1	45.72	45.48	45.36	45.68
Number of standard deviations away from the null hypothesis	$Z = \dfrac{n(p - p_0)}{\sigma}$	36.03	28.92	27.22	36.92	26.7	20.14	18.7	25.71

Of the n_0 points t^* are in a downtrend, a fraction p_0 satisfy $A1$, namely $P(t^* + 3) < P_{\text{avg}}(t^* + 4, t^* + 5, t^* + 6)$. If in addition there is a candlestick reversal pattern in points $(t^* + 1, t^* + 2, t^* + 3)$ then there are a total of n points of which a fraction p satisfy $A1$, deviating from the null hypothesis $(p = p_0)$ by 36.03 standard deviations. The analogous situation for an uptrend is described by B1 and 26.7 standard deviations is obtained. The conditions A2, A3, A4, and B2, B3, B4 are modifications of A1 and B1 respectively that establish robustness.

30. Since Z is proportional to $n^{1/2}$, this means that the values of Z would be reduced by a factor of $(30)^{1/2} = 5.5$. The, the Z value for ($A1$) is reduced from 36.03 to 6.58 while ($B1$) has its value reduced from 26.7 to 4.88, etc. Consequently, even with exaggerated assumptions on dependency, one has a high degree of confidence in the statistical significance of the result. In fact, since we are using the open and close of three days' trading, it is unlikely that a large number of stocks in one sector will all have a morning star, for example, during the same time period.

   The results show that the patterns provide an excellent short-term prediction for the course of prices. In fact, the TWS and the TIU patterns are predictive about three-fourths of the time for most of the entire data sets (Tables 4 and 5). (It is not surprising that the TOU and MS have somewhat less predictive power since the uptrend is established for just one and a half days and one day, respectively, in these two patterns, unlike the three days of the TWS and two days for the TIU.) If one looks at this from the perspective of a changing balance of supply and demand stock, then 75% endogenous predictive power leaves a small amount of room for stochastic exogenous events that will alter the balance of supply and demand. For example, a TWS pattern in a European country

Table 3. The average return per trade before costs. For the tests performed on the S&P 500, r_b shows the average profit (before costs such as commissions or the bid–ask spread) on trades indicated by the signals TWS, TIU, etc., as well as the average for all buy signals. The holding period is an average of two days. Similarly, r_s indicates the average profit on short sales when a (short) sale is triggered by each of the signals TBC, TID, etc. with the same holding period.

A1	11803/26386	B1	13672/24380
%	44.73	%	56.07
A2	14149/26386	B2	14001/24380
%	53.62	%	57.42
A3	14534/26386	B3	14260/24380
%	55.08	%	58.49
A4	11985/26386	B4	13444/24380
%	45.42	%	55.14

Table 4. Statistics with pattern.

	A1	*%*	*A2*	*%*	*A3*	*%*	*A4*	*%*
3WS	45/70	64.28	50/70	71.42	48/70	68.57	47/70	67.14
3IU	13/16	81.25	13/16	81.25	13/16	81.25	13/16	81.25
3OU	7/13	53.84	7/13	53.84	7/13	53.84	7/13	53.84
MS	23/41	56.09	24/41	58.53	23/41	56.09	23/41	56.09
Total	88/140	62.85	94/140	67.14	91/140	65	90/140	64.28

Table 5. Statistics with pattern.

	B1	%	B2	%	B3	%	B4	%
3BC	57/94	60.63	53/94	56.38	56/94	59.57	56/94	59.57
3ID	8/14	57.14	9/14	64.28	8/14	57.14	9/14	64.28
3OD	20/26	76.92	19/26	73.07	18/26	69.23	20/26	76.96
ES	103/146	70.54	100/146	68.49	105/146	71.91	105/146	71.91
Total	188/280	67.14	181/280	64.28	187/280	66.78	190/280	67.85

fund indicates that the balance of buying and selling has shifted in favour of the bulls, but an event such as a central bank rate hike would result in a shift toward more selling that is not reflected in the pattern. The amount of new selling would overwhelm the buying interest that was evident in the TWS pattern. A knowledgeable technician would not take on a long position under these conditions. This aspect of the limitations of any type of technical analysis should not be neglected in a practical implementation.

A second set of tests concerns the amount of profit per trade. We compute the profit or loss that would result from purchasing a stock at the close (as a buy signal appears in the candlestick pattern shortly before the trading day ends) and subsequently selling one-third of the stock on each of the following three days. The percentage profit from each trade is then given by:

$$r_b = \frac{\{P(t^* + 4) + P(t^* + 5) + P(t^* + 6)\}/3 - P(t^* + 3)}{P(t^* + 3)}$$

Similarly, we compute the profit or loss from a short-sale after a sell signal using the formula

$$r_s = \frac{P(t^* + 3) - \{P(t^* + 4) + P(t^* + 5) + P(t^* + 6)\}/3}{P(t^* + 3)}$$

For all four down-to-up reversal patterns the average rate of return, r_b, is found to be 0.9% for Data Set 2, which is highly significant in that each investment dollar is committed for an average of two days (which would result in annual compounding to 309% of the initial investment). In fact, there are an average of about five buy signals and eight sell signals per day, so there is ample opportunity to trade on this basis. For the up-to-down reversal patterns, the rate of return on short-sales, r_s, is 0.27% so that the initial investment is compounded annually to 140%. This is also highly significant since the S&P 500 was rising steadily throughout most of the time period studied. In particular, each of the patterns individually showed a return that is very significant compared with the null hypothesis of the average gain for an identical holding period, as shown in Table 3. The trading based on these candlestick patterns could be implemented in practice with moderate amounts of capital. The daily volume in the S&P 500 stocks is sufficiently large that one can place the trades at the close ('at market') without distorting the market. The largest cost is the bid–ask spread which is generally in the range of 0.1% to 0.3%. The commissions have been rather small at the deep discount brokers for decades and have recently become almost insignificant with electronic trading in recent years. A typical price is about $20 for several thousand shares. The profit per trade on the

buy side would then average between 0.56% and 0.76% after commissions and the bid–ask spread on a $100 000 trade. On a yearly basis each unit of capital would be compounded into 202% to 259% of the initial investment. One problem, however, may be increased volatility near these turning points that may result in a higher bid–ask spread. On the (short) sell side there would be limitations imposed by the uptick rule to the NYSE, but not on some other exchanges. Unless one is implementing trades with amounts that are orders of magnitude larger, one would not expect to alter the bid or ask price as trades are placed near the close, given the volume traded on the S&P 500 stocks.

Although computing the growth rate adjusted for costs provides an indication of the power of the method, technicians would use these methods in conjunction with other methods that would be aimed at increasing the profit per trade while the cost remains constant. In practice, one might use many other indicators in conjunction with these. These may include other types of intermediate (weeks) indicators for the particular stock as well as the overall market, interest rates and a particular commodity if the company uses or produces one. For example, in trading the stock of an oil-producing company, one would examine the patterns in prices and volume, both short and intermediate term, in the spot price of crude oil, the index of oil-producing companies, the overall US market and interest rates in addition to any fundamental analysis that affects oil supplies and the company. One would enter a trade on the buy side if the preponderance of these indicators, with a strong emphasis on the indicators of the company, were positive. Consequently, with the use of a number of rules, the costs become a smaller share of the profit per trade.

These tests on return rates confirm the predictive power of the patterns established above using nonparametric criteria, and exclude the theoretical possibility that a large number of profits are negated by a smaller number of larger losses.

From a game theoretic perspective, a trader who has the same information as others plus the knowledge of this method will have a competitive advantage. The patterns we have studied are a small fraction of those that are possible. If other patterns are equally predictive, the trader with this knowledge would consistently be able to place trades at more opportune times, while having the same costs as the others, and accumulate significantly greater profits.

4. Conclusions

The out-of-sample tests on two distinct sets of data provide a very high degree of certainty that the three-day patterns in candlestick analysis have predictive value. To the best of our knwowledge this is the first time a scientific test has shown statistical validity of any price pattern. We have devised a test that is almost entirely nonparametric by retaining key features of the patterns (involving inequalities between open and close) but eliminating considerations, such as the magnitude of the trading range, etc. This of course provides a slight bias against candlesticks, and the main result is strengthened.

The main conclusions can be summarized as follows.

(1) Our methodology differs from most large-scale statistical finance studies which examine data in a completely deductive manner, and complements them by testng a hypothesis that would be difficult to derive or formulate using purely statistical or neural network methods. The discussion of the large amount of data (Malkiel, 1995) on randomness in markets has taken on an interesting twist in recent years, as simple nonlinear deterministic models have been shown to be indistinguishable,

using linear statistical methods, from random systems (see Brock, 1986; Griliches and Intriligator, 1994 and references therein for more discussion). The patterns we test have been developed over centuries through trading experience. A statistical validation of these methods implies that experience in observing the behaviour of prices has a differential effect on the profitability of traders. The traditional argument has been that this type of experience could have no trading value since that information has already been incorporated into market prices. Consequently, the examination of patterns that offer basic clues to the mechanism of the market place are of fundamental value.

(2) The results of our study provide strong evidence that traders are influenced by price movements and probably use them as an indication of the positions of the other traders, particularly with respect to the changing balance of supply and demand.

(3) Our results can be interpreted in terms of the second derivative, $p''(t)$, which is considerably more difficult to detect than the first derivative. More specifically, if one expects a brief turn after a lengthy trend then the turning point will be difficult to detect due to its brevity. An examination of the TWS pattern, for example, indicates a clear negative derivative followed by a more brief positive derivative. The rationale for the particular pattern is that three consecutively higher closing prices alone would not be sufficiently decisive evidence of a trend reversal, and waiting longer for the establishment of an uptrend would imply missing a good part of the rebound. However, the Three White Soldiers pattern comprises three days of not only successively higher closing times, but positive movements throughout each day. The initial selling pressure is overwhelmed by further buying interest during the course of the day. The consistent set of six observations,

$$o_1 < o_2 < c_1 < o_3 < c_2 < c_3$$

provides enough information to indicate that the (moving average) downtrend has reversed and the bulls now have the upper hand. The discretized first and second derivatives with $h = 3$ provide a more analytic perspective:

$$f'(t) \approx \frac{f\left(t + \frac{h}{2}\right) - f\left(t - \frac{h}{2}\right)}{h}$$

$$f''(t) \approx \frac{f(t + h) - 2f(t) + f(t - h)}{h^2}$$

A nine-day period then is reduced to three periods of three days. Then f' is negative on the first two and positive on the last one for TWS. The second derivative is insignificant during the first six-day period and positive on the latter, thereby indicating a reversal from downtrend to uptrend.

(4) Another aspect of our study concerns the profitability of the method by examining the gain or loss on each trade compared with the null hypothesis of the average gain for an identical holding period. The results were significant for both the buy and the sell signals, with the buy signal resulting in a tripling of the initial investment during a one-year period (with costs taken into account). While this result provides a strong confirmation, one could presumably obtain even stronger practical results by using parametrization and additonal methods as discussed below.

(i) The length of the candles and the magnitude of the downtrend are important indicators of the decisiveness of the pattern. Also, the cost of the transaction may force an additional filter that

involves parametrization. Consequently, a practical use of these methods could be enhanced by determining these parameters to provide a more restrictive but more profitable set of trading opportunities.

The patterns we study here are only one of dozens of indicators a technician might use, usually in a complex combination. For example, the candlestick patterns may be used to pinpoint a bottom that has been developing for several weeks in the form of an inverted head and shoulders pattern. Hence a practical application could be enhanced by utilizing candlesticks in conjunction with slightly longer time scale methods.

(ii) From the perspective of objective study of the nature of markets, there are additional reasons for using essentially nonparametric tests. One is that a small number of stocks cannot distort the statistics. For example, Table 2 shows that of the 4688 points in a downtrend with a candlestick pattern, the observed number of reversals (i.e. ($A1$) is satisfied) of 3339 would need to drop to $2111 + 72 = 2183$ before 95% confidence is lost. Whether or not a handful of stocks provide much more profit as a result of these methods, the validity of the methodology for a broad class of stocks cannot be negated based on considerations involving a small group of stocks.

From an economics and finance perspective the significance is not so much the existence of a set of patterns that have a tremendous predictive power, but rather that the underlying assumptions in this study of markets need to be re-examined.

While the statistical validity of a set of price patterns does not, in itself, offer a replacement for some of these basic assumptions, it is difficult to avoid the conclusion that evidence of a turnaround in prices tends to yield higher prices. This means that traders are reacting to the expectations involving the strategies and resources of the other participants. Thus, our study can be seen as lending support to a game theoretic approach involving imperfect information (about value) and a finite amount of resources and arbitrage.

From a mathematical microeconomics perspective, the presence of a second derivative (discussed in item (3) above) that has deterministic origins is incompatible with the modern price theory approach (as discussed in standard texts such as Watson and Getz, 1981) in which the price derivative depends only upon price. The existence of deterministic oscillations is mathematically equivalent to the presence of a second derivative in the price equation arising, for example, from (i) the dependence of supply and demand on the price of previous time period; and (ii) the dependence of supply on the price derivative.

References

Beard, T. and Beil, R. (1994) Do people rely on the self-interested maximization of others? An experimental test, *Management Science*, **40**, 252–262.

Blume, L., Easley, D. and O'Hara, M. (1994) Market statistics and technical analysis: the role of volume, *Journal of Finance*, **69**, 153–181.

Brock, W.A. (1986) Distinguishing random deterministic systems, *Journal of Economic Theory*, **40**, 168–195.

Brock, W.A., Josef, L. and LeBaron, B. (1992) Simple technical trading rules and stochastic properties of stock returns, *Journal of Finance*, **47**, 1731–1764.

Caginalp, G. and Balenovich, D. (1996) Trend based asset flow in technical analysis and securities marketing, *Psychology and Marketing*, **13**, 401–444.

Caginalp, G. and Constantine, G. (1995) Statistical influence and modeling of momentum in stock prices, *Applied Mathematical Finance*, **2**, 225–242.

Griliches, Z. and Intriligator, M. (eds) (1994) *Handbook of Econometrics IV* North Holland, Amsterdam.

Kamijo, K. and Tanigawa, T. (1993) Stock price pattern recognition: a recurrent neural network approach, in *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real World Performance*, R.R. Trippi and E. Turban (eds), Probus, Chicago, pp. 357–369.

Lo, A. and MacKinley, A.C. (1988) Stock market prices do not follow random walks: evidence from a simple specification test, *Review of Financial Studies*.

Malkiel, B.G. (1995) *A Random Walk Down Wall Street* (6th edn), Norton, New York.

Meyerson, R. (1978) Refinements of the Nash equilibrium concept, *Int J Game Theory*, **7**, 73–80.

Morris, G.L. (1992) *Candlepower, Advanced Candlestick Pattern Recognition and Filtering Techniques for Trading Stocks and Futures*, Probus, Chicago.

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

Pistolese, C. (1994) *Using Technical Analysis, a Step-by-step Guide to Understanding and Applying Stock Market Charting Techniques*, Irwin, Chicago.

Porter, D. and Smith, V. (1994) Stock market bubbles in the laboratory, *Applied Mathematical Finance*, **1**, 111–128.

Rosenthal, R. (1981) Games of perfect information, predatory pricing and the chain store paradox, *Journal of Economic Theory*, **24**, 92–100.

Shiller, R.J. (1981) The use of volatility measures in assessing market efficiency, *Journal of Finance*, **36**, 291–304.

Watson, D.S. and Getz, M. (1981) *Price Theory and Its Uses*, 5th edn, University Press of America, Lanham, Maryland.

White, H. (1993) Economic prediction using neural networks: The case of IBM daily stock returns, in *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real World Performance* R.R. Trippi and E. Turban (eds), Probus, Chicago, pp. 315–328.