Modelling for Forecasting of Pattern Recognition

- Based on comparison and analysis be-tween U.S. stock Market and Chinese stock Markets

Hong Zhou[1,*], Yu He[2,3] and Yuxiang Jin[3]

[1]School of Finance, Renmin University of China, Beijing, 100872, China
[2]School of Finance, Zhejiang University of Finance & Economics, Hangzhou, 310018, China
[3]China Academy of Financial Research, Zhejiang University of Finance & Economics, Hangzhou 310018, China
[4]Galaxy Derivatives Financial Service Co.Ltd, Beijing, 100045, China
*Corresponding author

Abstract—**An essential aspect of stock trading is the accurate forecast of stock price. This enables buy and sell points to be determined, which facilities profitability whilst reducing potential loses. This paper proposes a "Two-stage pattern Strategy (TSPS)" as an effective and intuitive mechanism to identify buy and sell signals.**

Utilizing technical analysis methods and pattern recognition principles, the TSPS indicates that (1) the trend to be more important than the isolated price; (2) two continuous unidirectional trends could verify the uptrend or downtrend; (3) the high or low price have more prediction power than the closing price; (4) the low price is more effective in prediction in the uptrend case, while the high price is more valuable than the low price in the downtrend case.

Accordingly, this paper establishes mechanism to recognize up or down patterns and the pinpoint for buying or selling. This empirical study was done to verify the prediction power and profitability of TSPS. This study compared the performance of the U.S. and Chinese markets. The results show that TSPS can be widely used in the market regardless of the economic environment

Keywords-trend forecasting; pattern recognition; U.S. and Chinese stock markets

I. INTRODUCTION

The discovery of price trends as well as the timing of buy and sell points are key components in stock market prediction. Early theories that were used to conduct technical analysis to forecast price fluctuations included Dow's theory (Hamilton 1922), Elliott wave theory (Elliott, 1994) and Gann's theory. However, both Dow's theory and Elliott wave theory are rather ambiguous with respect to quantitative analysis. With the rapid advance in computer processing power, computer-assisted techniques have been used to predict stock prices through the use of computer software, algorithms and mathematical models such as the Kalman filter method and artificial neural networks method. Our paper want to discover a way to determine the trend and to identify the buy or sell signal remains a topic worthy of additional study.

Some scholars try to define a trend using chart such as candlestick charting, such as Zapranis, Tsinalanidis[1], Park[2], and Bigalow[3]. Caginalp and Laurent[4] show in great detail how to define trends and recognize candlestick patterns. K.-i. Kainijo and T. Tanigawa [5] employed five-day moving average to identify uptrends/downtrends. They assert that upward trend is defined as the increase in average price whilst a decrease indicates a downward trend.

Along with a greater focus on candlestick analysis, new approaches that recognize and identify candlestick chart patterns have been developed. Y.-M. Shiu and T.-H. Lu [6], T.-H. Lu, et al. [7], investigated the profitability of two-day candlestick patterns by purchasing stock based on bullish (bearish) patterns and holding until bearish (bullish) patterns occurred. They found three bullish reversal patterns are profitable in the Taiwan stock market. C.-H. L. Lee, et al. [8], proposed an approach which would extract fuzzy candlestick patterns from a financial time series and identify a set of patterns to aid the investment decision making process. Alternatively, A. Zapranis and P. E. Tsinaslanidis [1] also identified a rigorous rule-based mechanism for identifying the rounding bottoms (also known as saucers) pattern and resistant levels.

As we know, investors benefit more from timely identification and use of the turning point, a novel approach for identifying the pinpoints for buying and selling is required. This paper proposes a two stage pattern Strategy (TSPS) method which identifies the trends and timing turning points. It further showed that the high price or low price of the candlestick can be more important than the closing price in forecasting the turning trend. When a downtrend commences, the high price conveys more information than the low price. Conversely, in an uptrend the low price is more informative than the high price.\

II. MODELING

Analyzing a stock's price history could reveals vital information that can be used to enable timely and accurate predictions. The TSPS method utilises chart to identify the change of the price of a stock. We use different chart pattern to define a different price movement trend.

As usual, we define the chart bar is comprised of four prices in a specific time period: the high price, low price, opening and closing price. The range between the opening and closing prices is referred to as the real body, illustrated by a rectangle box.
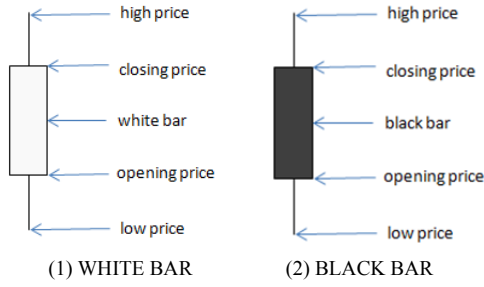
(1) WHITE BAR　　(2) BLACK BAR

FIGURE I. CHART PATTERN DEFINITION

When the closing price is higher than opening price, it is called a white bar; subsequently, if it is below, it is called a black bar. The price excursions above and below the real body are, respectively, called the upper and lower shadows. The white bar indicates an uptrend, while the black bar indicates an downtrend. (See Fig. 1)

The TSPS method proposes that the emergence of two consecutive white bars or two black bars indicates either an uptrend or downtrend could be affirmed. Regardless of multiple candlesticks with identical color or the depth of the upper and lower lines, it is two consecutive opposite color candlesticks that indicate the emergence of the two candlestick pattern.

Assumption:

(1) Event A denotes the emergence of the reversal, and all the graphic patterns constitute population Ω;

(2) Bi denotes a graphic pattern that is comprised solely of either black or white bar. Furthermore, i represents the number of bars in this pattern. For example B_1 denotes one black bar after one or more white bars. It can also denote one white bar after one or more black bars; B_2 denotes two consecutive same colored bars; B_N denotes the maximum amount of continuous single colored bars.

(3) The greater the number of single color bar, the higher the probability of predicting an uptrend or downtrend.

Regardless of the graphic pattern's appearance, it can be easily classified as Bi and get:

$$p(B_i \mid A) = \frac{P(B_i)P(A \mid B_i)}{\sum_{i=1}^{n} P(B_i)P(A \mid B_i)}$$

Calculation of P(Bi/A), (i=1,2,…,N) will identify the pattern with high probability of detecting the trend reversal.

III. ALGORITHM AND STRATEGY

In order to identify the most probable and timely pattern that may indicate a trend reversal, each pattern must be converted from a pictorial representation into a written form and expressed numerically.

To represent a trend reverse, a comparison of two or more prices was used. By utilizing either the high or low price from two or more bars, an upward or downward trend is indicated. We create two Strategy to prove it .

Strategy I: When a white bar appears after one or more black bars, the upturn is indicated. The high price and low price, when compared with the next candlestick's corresponding price, will be lower indicating an uptrend. Conversely, when a black bar appear after one or more white bars a downturn has indicated. If the two prices are higher than the later candlestick, a downtrend is indicated.

Strategy II; we add an additional condition to further illustrate when a trend is formed. It is hypothesized that when a trend reversal occurs, the low or high price becomes a better predictor of future movement. In an uptrend, the low price becomes the better indicator. Conversely, the high price becomes a better indicator in a downtrend. To increase the predictive power, the low price is compared with the corresponding price of the previous candlestick in an uptrend. Alternatively, the high price is used in a downtrend. Once an upturn or downturn is confirmed, the prices are then compared with the following candlestick.

IV. EMPIRICAL ANALYSIS

A. Data Selection

Shanghai and Shenzhen Stock Exchanges are China's main stock markets and NASDAQ stock exchanges is the main stock market of American. These three markets represents the emerging market and mature market respectively. This paper selected data from these three markets.

Shanghai and Shenzhen trading records from Jan. 1st 2006 to Dec. 31st 2008 were chosen as the study samples. Whilst this period conclude a whole cycle from bullish to bearish. Market. During the sample period, 867 listed companies have recorded trades in Shanghai Stock Exchange and 754 companies in Shenzhen Stock Exchange. As the peak point presents in Oct. 16th 2007, two parts of the period are divided to verify that Strategy I & II fit regardless of bullish market or bearish market.

This paper randomly selected a single years trading records in mature capital market as the research period. Daily prices between Jan. 1st 2010 to Dec. 31st 2010 in the NASDAQ market was chosen as the study sample. Accordingly, there were 4460 listed companies in the NASDAQ during our sample period.

B. Strategy Accomplish

Set H_1, L_1, O_1 and C_1 indicate the high, low, opening and closing price of the first day of the pattern, and H_2, L_2, O_2 and C_2 refer to the high, low, opening and closing price of the second day of the pattern. Suppose H_3, L_3, O_3 and C_3 indicate the high, low, opening and closing price of the first day after the pattern.

Let r, w, N refer to the trend forecasting right times and wrong times, and the total days in the sample interval, and H(i), L(i), O(i), C(i) indicate the high, low, open and closing price of the day i, $i \in [1,N]$.

The accurate rate indicates the effective of the TSPS.

For the Strategy I , the algorithm as follow:

Set r = 0, w = 0, H_1 = H(1), H_2 = H(2), H_3 = H(3), L_1 = L(1), L_2 = L(2), L_3 = L(3), O_1 = O(1), O_2 = O(2), O_3 = O(3), C_1 = C(1), C_2 = C(2), C_3 = C(3)

For i = 3: N
    If $O_1 > C_1$, and $O_2 < C_2$, then
        If $L_3 > L_2$, then
            r = r + 1
        Else
            w = w + 1
        end if
    end if
    H_1 = H_2, H_2 = H_3, L_1 = L_2, L_2 = L_3,
    O_1 = O_2, O_2 = O_3, C_1 = C_2, C_2 = C_3,
    i = i + 1
    H_3 = H(i), L_3 = L(i),
    O_3 = O(i), C_3 = C(i)
Loop
Accurate rate = r/(r+w)

For the Strategy II , the algorithm as follow:

Set r = 0, w = 0, H_1 = H(1), H_2 = H(2), H_3 = H(3), L_1 = L(1), L_2 = L(2), L_3 = L(3), O_1 = O(1), O_2 = O(2), O_3 = O(3), C_1 = C(1), C_2 = C(2), C_3 = C(3)

For i = 3: N
    If $O_1 > C_1$, $O_2 < C_2$, $L_2 > L_1$, then
        If $L_3 > L_2$, then
            r = r + 1
        Else
            w = w + 1
        end if
    end if
    H_1 = H_2, H_2 = H_3, L_1 = L_2, L_2 = L_3,
    O_1 = O_2, O_2 = O_3, C_1 = C_2, C_2 = C_3,
    i = i + 1
    H_3 = H(i), L_3 = L(i),
    O_3 = O(i), C_3 = C(i)
Loop
Accurate rate = r/(r+w)

C. Result of Accuracy rate

The Accuracy rate result of TSPS in different markets and periods are shown in Table 1 and Table 2. This paper calculates the distribution of patterns in the time of reserve, and find the P (B2/A) is more than 60%. The result prove that the two stage pattern is effect.

TABLE I. PREDICTION POWER OF "STRATEGY I" IN DIFFERENT MARKETS AS WELL AS BULLISH AND BEARISH MARKET

	Companies	Black to white bar			white to black bar		
		Total times	*high price*	*low price*	*Total times*	*high price*	*low price*
Shanghai Stock Exchange (Jan 1st. 2006~Dec 31st. 2008)	867	139685	71.60%	76.77%	140049	74.22%	73.02%
Shanghai Stock Exchange Bullish market (Jan 1st, 2006 ~ Oct 16th, 2007)	853	79938	73.26%	78.70%	80399	70.93%	68.82%
Shanghai Stock Exchange Bearish market (Oct 17th, 2007~Dec 31th, 2008)	853	59747	69.38%	74.20%	59650	78.67%	78.69%
Shenzhen Stock Exchange(Jan. 1st 2006~Dec. 31st 2008)	754	100965	71.28%	76.93%	101127	75%	73.75%
Shenzhen Stock Exchange Bullish market (Jan 1st, 2006 ~ Oct 16th, 2007)	655	53210	72.56%	79.06%	53492	71.57%	69.50%
Shenzhen Stock Exchange Bearish market (Oct 17th, 2007 ~ Dec 31th, 2008)	722	47755	69.85%	74.55%	47635	78.85%	78.53%
NASDAQ Stock Market (Jan. 1st 2010 ~ Dec. 31th 2010)	Total	120637	67.46%	68.34%	109932	65.91%	65.07%

TABLE II. PREDICTION POWER OF STRATEGY II" IN DIFFERENT MARKETS AS WELL AS BULLISH AND BEARISH MARKET

	Companies	Black to white bar			white to black bar		
		Total times	*high price*	*low price*	*Total times*	*high price*	*low price*
Shanghai Stock Exchange (Jan 1st, 2006~Dec 31st, 2008)	867	61274	71.35%	71.60%	72895	72.21%	71.06%
Shanghai Stock Exchange Bullish market (Jan 1st, 2006 ~ Oct 16th, 2007)	853	38278	72.15%	74.09%	41138	69.17%	67.19%
Shanghai Stock Exchange Bearish market (Oct 17th,	853	22996	70.02%	67.47%	31757	76.14%	76.06%

2007~Dec 31th, 2008)							
Shenzhen Stock Exchange(Jan. 1st 2006~Dec. 31st 2008)	754	43570	70.14%	71.10%	53627	73.19%	71.73%
Shenzhen Stock Exchange Bullish market (Jan 1st, 2006 ~ Oct 16th, 2007)	655	24900	70.62%	74.14%	27799	70.09%	67.94%
Shenzhen Stock Exchange Bearish market (Oct 17th, 2007 ~ Dec 31th, 2008)	722	18670	69.51%	67.03%	25828	76.53%	75.81%
NASDAQ Stock Market (Jan. 1st 2010 ~ Dec. 31th 2010)	4460	234456	67.12%	71.09%	233445	67.34%	63.19%

1) Accuracy rate of Strategy I

Shown as Table 1, in the Shanghai stock exchange transactions, of 139,685 times that black bars changed to white bars, Strategy I's accuracy rate was 71.60% of the high price. That is to say, 71.60% of the high price in a white bar was higher than that in the later bar. Compared with the low price, the accuracy rates was 76.77% . While in Shenzhen stock market, the accuracy rate according to the high, low in the case of black bar to white bar was 71.28%, 76.93% respectively; in NASDAQ market, they were 67.46%, 68.34%.

The above results verify Strategy I is effective.

2) Accuracy rate of Strategy II

As shown in Table 2, in the Shanghai stock exchange transactions, of 139,685 times that black bars changed to white bars, Strategy II's accuracy rate was 71.35% of the high price. That is to say, 71.35% of the high price in a white bar was higher than that in the later bar. Compared with the low price, the accuracy rates was 71.60% . While in Shenzhen stock market, the accuracy rate according to the high, low price in the case of black bar to white bar was 70.14%, 71.10%; in NASDAQ market, they were 67.12%, 71.09%.

The above results verify Strategy II is effective.

3) Comparison result with other popular indicators

To examine TSPS deeply, we establish two corresponding trade strategies to execute imitation trading. And we compared with RSI, MACD and KDJ strategies, which are popular indicators used by investors. Yield results are list in Table 3.

TABLE III. COMPARISON WITH OTHER POPULAR INDICATORS

Strategy	average yield	average yield	average yield
	NASDAQ market	Shanghai Exchange	shenzhen Exchange
Strategy I	9.00%	19954.2%	8148.1%
Strategy II	15.18%	22482.1%	13570.6%
RSI	17.50%	0	2.2%
MACD	-2.68%	4.7%	0.9%
KDJ	1.54%	0	0

From the result the average return based on either Strategy I or Strategy II yields a good profit margin. Compared with other indicators, such as RSI, MACD, KDJ, our strategies are effective.

V. CONCLUSION

This paper propose a two stage pattern Strategy as a timely and predictable pattern that indicate stock price trend reverse. A model is established to rigorously support the proving progress of the effective TSPS method.

The studies have found that the emergence of two same colored lines after continuous opposite colored lines indicates a reverse in the stock trend. That is, the emergence of two opposite continuous bars can indicate a trend reversal, a strategy known as the "Two-stage Strategy" or "Two-line Strategy", which referred as TSPS.

REFERENCES

[1] A. Zapranis and P. E. Tsinaslanidis, "A novel, rule-based technical pattern identification mechanism: Identifying and evaluating saucers and resistant levels in the US stock market," Expert Systems with Applications, vol. 39, pp. 6301-6308, 2012.

[2] C.-H. Park and S. H. Irwin, "What do we know about the profitability of technical analysis?," Journal of Economic Surveys, vol. 21, pp. 786-826, 2007.

[3] S. W. Bigalow, Profitable candlestick trading: pinpointing market opportunities to maximize profits vol. 500: John Wiley & Sons, 2011.

[4] Caginalp, G. and H. Laurent (1998). "The predictive power of price patterns." Applied Mathematical Finance 5: 181-206.

[5] K.-i. Kainijo and T. Tanigawa, "Stock Price Pattern Recognition - A Recurrent Neural Network Approach," pp. 215-221, 1990.[6] Y.-M. Shiu and T.-H. Lu, "Pinpoint and synergistic trading strategies of candlesticks," International Journal of Economics and Finance, vol. 3, pp. 234-244, Feb. 2011.

[6] T.-H. Lu, Y.-M. Shiu, and T.-C. Liu, "Profitable candlestick trading strategies—The evidence from a new perspective," Review of Financial Economics, vol. 21, pp. 63-68, 2012

[7] C.-H. L. Lee, Y.-C. Liaw, and L. Hsu, "Investment decision making by using fuzzy candlestick pattern and genetic algorithm," presented at the Fuzzy Systems (FUZZ), IEEE International Conference on Date 27-30 June 2011.

[8] C.-H. L. Lee, A. Liu, and W.-S. Chen, "Pattern discovery of fuzzy time series for financial prediction," presented at the Knowledge and Data Engineering, IEEE Transactions on Vol.18, Issue: 5, 2006, pp. 613-625.

.