

Candlestick Analysis in Forecasting U.S. Stock Market: Are They Informative and Effective

Haixuan Qiu

Central China Normal University
No.152 Luoyu Road, Hongshan District,
Wuhan, Hubei, China
e-mail: 156374615@qq.com

Fanzhuoqu Liu

Southwestern University of Finance and Economics
No.555 Liutai Road, Wenjiang District,
Chengdu, Sichuan, China
e-mail: 717026823@qq.com

Abstract— Stock price prediction is one of the hottest topics of research in both academia and industry. Being able to predict the trend of price correctly allows investors to gain profit. There have been multiple strategies in stock price prediction, such as multiple machine learning methodologies and forecast from sentiment analysis of the public and news feedings. Among these strategies, one of the oldest but still widely used strategy is the candlestick analysis, which is a simple way that allows general investors to predict the market trend. However, there lacks an unbiased estimation of the effectiveness of such a method. In this paper, using an unbiased and rigorous way to test multiple U.S. stocks, we were able to show that most of the candlestick patterns were not informative, while a small fraction of them provides some correct information for market trend compared with random guesses. Our study serves as a stepping-stone for re-evaluating the candlestick analysis and urges more similar and thorough studies to be conducted to guide the general public in stock market investment better.

Keywords- trading strategy; technical analysis; candlestick; stock price

I. INTRODUCTION

Stock price prediction has been one of the hottest research topics in recent years. The capability to predict the stock price comparatively accurately allows investors to gain a considerable amount of profit. The stock market could be highly fluctuating, ranging from the long-term economic boom and bust cycles [1] that can last for several years or even decades to fast fluctuation at the scale of millisecond that leads to an arms race in the field of the high-frequency trading [2]. As a result, there are also many different ways of technical analysis and various strategies for trading. Among them, the simplest one is the buy-and-hold strategy, for which there is no need for investors to watch stock price closely, and the profit comes from the overall increasing trend of the selected portfolio [3]. This trading strategy was highly recommended by many investors, the most famous one being Warren Buffett. At the other end of the extreme is the area of the high-frequency trading performed by several large trading firms using secretive proprietary trading strategies [4].

Apart from these extremes at scale for stock price changes, general investors are mostly interested in daily changes of stock prices. High-throughput data mining with the help of computers is widely used in various areas of

research, such as bioscience [5, 6] and engineering [7, 8]. Not surprisingly, machine learning algorithms are also widely used for stock price prediction [9]. For instance, decision tree and support vector machines were commonly used for stock price prediction [10, 11].

However, for the general public that lacks in-depth knowledge of machine learning, many of them prefer candlestick analysis, which is one of the oldest but still widely used strategy. Unlike other methodologies that call for statistical and computational background, candlestick analysis is a simple and straightforward method to predict the stock price. Specifically, individual investors, by reading the K-chart, make personal judgments on whether or not the K-chart fits one of the candlestick patterns, and make decisions based on the predicted market trend to buy or sell. Candlestick analysis assumes that the high, low, open and close prices are information sufficient and provide a visually appealing way for people to look at the trend of the stock price changes. It has been widely used by many major financial websites and analysis software [12].

However, there lacks general information on the validity and effectiveness of the candlestick analysis, as many textbooks and previous publications are teemed with individual cases that substantiate the correctness of such patterns, leading to a possible illusion that the prediction of the candlestick analysis is supposed to be correct, at least to a large extent. In this paper, we studied six very popular candlestick patterns and calculated their correctness using multiple US stocks. Our conclusion shows that many of the patterns did not have a percentage of correct prediction higher than 50%. However, we did show that some patterns were significantly different from random guesses. Our study calls for a re-evaluation of the candlestick models and report of the correctness for each patterns and urges the public to realize the potential risk and return before extensively applying such technical analysis method as their major guideline for trading.

II. METHODS

A. Downloading Stock Price Data

We downloaded a series of US stocks, with the name of the stock and companies listed as in Table I. Although candlestick analysis is a comparatively old method, we are interested in evaluating its effectiveness in modern times. The full length of each stock was used for analysis.

TABLE I. LIST OF STOCK AND COMPANIES USED IN THIS STUDY

Stock	Company
FB	Facebook
MRK	Merck & Co.
MSFT	Microsoft Corp.
NDAQ	Nasdaq, Inc.
ORCL	Oracle Corp.
SPY	SPDR S&P 500 ETF
TGT	Target Corp.
TQQQ	ProShares UltraPro QQQ
UDR	UDR Inc

B. Describing Candlestick Patterns

Unlike many modeling methods that are highly quantitative, candlestick patterns allow each investor to make their own judgment on, for instance, what an uptrend and downtrend is. Here, we define an uptrend as the stock price in at least four out of five days to be increasing. As for the definition of a large body and small body, to allow for comparison across various stocks, we define large body to be within the top 50% among the past 30 days, and small body to be within the bottom 50% among the past 30 days. We expect small changes in the criteria will not change the general conclusions reached here. We varied these two criteria, and the effectiveness of each method is largely unchanged (Data not shown). A larger sample size will be available once we use looser criteria and the sample size will be smaller if we make the criteria more stringent.

The characteristics of each candlestick patterns were listed in Table II [13].

TABLE II. CANDLESTICK PATTERNS

Pattern	Prediction	Definition
Hanging man	Decrease	Considered a bearish pattern during an uptrend. Appears after an uptrend. A black or a white candlestick that consists of a small body near the high with a little or no upper shadow and a long lower tail. The next day confirms the trend of the downtrend.
Hammer	Increase	Considered a bullish pattern during a downtrend. Appears after a downtrend Same shape as the hanging man The next day confirms the uptrend
Dark cloud cover	Decrease	Considered a bearish pattern during an uptrend. Appears after an uptrend. Consists of a long white candlestick followed by a black candlestick that opens above the high of the white candlestick and closes well into the body of the white candlestick.
Piercing Line	Increase	Considered a bullish pattern during a downtrend. Appears after a downtrend. Consists of a black candlestick followed by a white candlestick that opens lower than the low of preceding but closes more than halfway into black body candlestick.

Evening Star		Considered a bearish pattern during an uptrend. Appears after an uptrend. Consists of a large white body candlestick followed by a small body candlestick (black or white) that gaps above the previous. The third is a black body candlestick that closes well within the large white body.
Morning Star		Considered a bullish pattern during a downtrend. Appears after a downtrend. Consists of a large black body candlestick followed by a small body (black or white) that occurred below the large black body candlestick. On the following day, a third white body candlestick is formed that closed well into the black body candlestick.

C. Searching for Candlestick Patterns

We scanned through the aforementioned nine stocks for the six candlestick patterns. For each pattern, we evaluate the price of the next day compared with 5 days later to determine the market trend. If the former is higher, the market is going down. If the former is lower, the market is going up. We then compared if the predicted trend is consistent with the actual trend and recorded the prices changes.

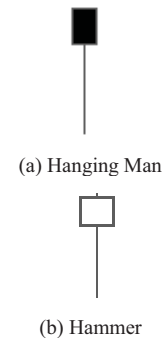
D. Quantifying the Effectiveness of Each Candlestick Patterns

For each candlestick patterns, we quantify its effectiveness from two aspects: percentage of correct prediction and expected revenue. For the former, we calculated the total number of cases that match the candlestick pattern (the column "Total number" in Table 3) and the number that predicts the market trend correctly. Using a binomial test, we can see whether the percentage of correct prediction was significantly different from 0.5 or not.

Moreover, for each case, we calculated the price differences between the closing price of the next day and five days later. The price difference represents the expected revenue in the short term. We calculated the median of the revenue in each case and used a Mann-Whitney U test to see if the expected revenue was significantly different from 0.

III. RESULTS

Simple illustrations (ignoring the uptrend and downtrend) for each pattern were shown in Fig. 1.



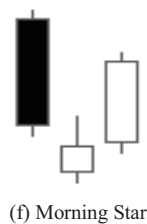
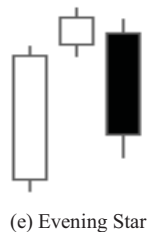
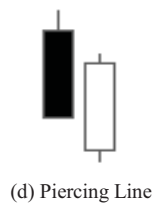
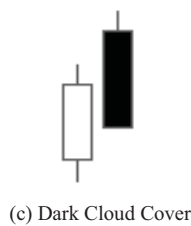


Figure 1. Simple illustrations of candlestick patterns studied. When the close price is lower than the open price, the body of the candle is black (such as (a)), while if the close is higher than the open, the body of the candle is white (such as (b)).

The percentage of correct guesses and the corresponding P-values calculated from a binomial test were shown in Table III.

TABLE III. . PERCENTAGE OF CORRECT PREDICTIONS FOR CANDLESTICK ANALYSIS

Pattern	Number of cases	Percentage of correct prediction	P—value from a binomial test
Hanging man	416	0.53	0.1516
Hammer	1118	0.56	0.0001
Dark cloud cover	48	0.52	0.4427
Piercing line	37	0.70	0.0100
Evening Star	37	0.46	0.7443
Morning star	23	0.61	0.2024

Among the six patterns, four of them were not significantly different from random guesses, while two of them provide some information on the market trend. The one

with the highest percentage of correct guesses is the piercing line, with 70% of correct guesses for a trend reversal. The second strongest one is the Hammer, which has 56% of correct guesses with a very significant p-value of 0.0001.

The median revenue for each candlestick pattern was listed in Table IV. Consistent with the percentage of correct predictions listed in Table III, among the six patterns, hammer and piercing line are the two patterns with the strongest signal for a profitable return. The other four patterns are not significantly different from 0.

TABLE IV. THE EXPECTED REVENUE FOR PATTERNS OF THE CANDLESTICK ANALYSIS

Pattern	Median revenue	P value calculated from the Mann-Whitney test
Hanging man	-0.0006	0.8845
Hammer	0.0078	4.0xE-10
Dark cloud cover	-0.0030	0.6302
Piercing line	0.0210	0.0026
Evening Star	-0.0065	0.2761
Morning star	0.0094	0.4634

In conclusion, among the six candlestick models, only two of them are somewhat effective and provides profitable revenue for investors using nine US stocks. Our results provide an up-to-date view of the “ancient” candlestick analyses. Given the limited applicability, general investors should not be using the candlestick model as the major guideline for their decision-making process.

Interestingly, it is worth noting that the Hammer and Piercing line were both reversal signals for bullish pattern during a downtrend, while none of the three reversal signals for bearish patterns were significant. Moreover, each of the reversal signal for uptrend (Hammer, Piercing line and Morning star) provide a positive median revenue, while each of the reversal signal for a downtrend (Hanging man, Dark cloud cover and Evening star) all gives a negative revenue. Given that many of the stocks in the list in the history were an uptrend, we would expect that there is a higher fraction of bullish reversal than bearish reversal for these stocks among the selected stocks and the expected revenue is positive.

IV. DISCUSSION

Through unbiased testing of candlestick patterns using nine US stocks, we were able to show that two out of the six candlestick patterns were effective and informative, while the other four are not. Therefore, before directly applying the candlestick analysis model to modern stock market data, investors should have a clear idea about their general correctness and the expected revenue.

There are some limitations to this study. Firstly, as pointed out in the result part, the overall trend of the stock market is going upward, therefore, there is a tendency for all bullish patterns to be correct and generate positive returns, and for all bearish patterns to be incorrect and generate negative returns. Here, we used a null expectation of 50% for the expectation of random guess and 0 for expected return given no signal, which might not be the most rigorous way. Secondly, all six patterns here all indicates trend reversal, while some trend could be long lasting. Therefore, we would

expect each of the candlestick patterns, if totally uninformative, to have a probability of correct prediction to be lower than 50%. After all, these warning signals signify the possibility that there may be an upcoming reversal of the market trend, instead of asserting that there must be a reversal of the trend. It would be interesting for further studies to address these two aspects.

As for general investors, it is worthwhile to note that, although there are many salient cases listed in previous publications and textbooks, associating these patterns with the expected outcome may not be wise given our results listed here. Even for the strongest pattern, the percentage of correct prediction is 70%, let alone the weaker patterns. Our study serves as a stepping-stone for future publications and textbooks to report the effectiveness and applicability of each pattern in a rigorous manner, which will undoubtedly benefit the investors a lot in decision making.

REFERENCES

- [1] Adam, Klaus, Albert Marcet, and Johannes Beutel. "Stock price booms and expected capital gains." *American Economic Review* 107, no. 8 (2017): 2352-2408.
- [2] Budish, Eric, Peter Cramton, and John Shim. "The high-frequency trading arms race: Frequent batch auctions as a market design response." *The Quarterly Journal of Economics* 130, no. 4 (2015): 1547-1621.
- [3] Maguire, Phil, Stephen Kelly, Robert Miller, Philippe Moser, Philip Hyland, and Rebecca Maguire. "Further evidence in support of a low-volatility anomaly: Optimizing buy-and-hold portfolios by minimizing historical aggregate volatility." *Journal of Asset Management* 18, no. 4 (2017): 326-339.
- [4] Boehmer, Ekkehart, Dan Li, and Gideon Saar. "The competitive landscape of high-frequency trading firms." *The Review of Financial Studies* 31, no. 6 (2018): 2227-2276.
- [5] Libbrecht, Maxwell W., and William Stafford Noble. "Machine learning applications in genetics and genomics." *Nature Reviews Genetics* 16, no. 6 (2015): 321.
- [6] Li, Chuan, Zhi Wang, and Jianzhi Zhang. "Toward Genome-Wide Identification of Bateson–Dobzhansky–Muller Incompatibilities in Yeast: A Simulation Study." *Genome biology and evolution* 5, no. 7 (2013): 1261-1272.
- [7] Wang, Xiaoyu, and Amir Mortazawi. "Bandwidth Enhancement of RF Resonators Using Duffing Nonlinear Resonance for Wireless Power Applications." *IEEE Transactions on Microwave Theory and Techniques* 64, no. 11 (2016): 3695-3702.
- [8] Cao, Song, Kan Chen, and Ram Nevatia. "Activity recognition and prediction with pose based discriminative patch model." In *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*, pp. 1-9. IEEE, 2016.
- [9] Patel, Jigar, Sahil Shah, Priyank Thakkar, and K. Kotecha. "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques." *Expert Systems with Applications* 42, no. 1 (2015): 259-268.
- [10] Basti, Eyup, Cemil Kuzey, and Dursun Delen. "Analyzing initial public offerings' short-term performance using decision trees and SVMs." *Decision Support Systems* 73 (2015): 15-27.
- [11] Fenghua, W. E. N., X. I. A. O. Jihong, H. E. Zhifang, and G. O. N. G. Xu. "Stock price prediction based on SSA and SVM." *Procedia Computer Science* 31 (2014): 625-631.
- [12] Nison, Steve. *The Candlestick Course*. Vol. 163. John Wiley & Sons, 2004.
- [13] Wikipedia: https://en.wikipedia.org/wiki/Candlestick_pattern .