Abstract

The report investigates the effectiveness of bullish reversal candlestick patterns. The study is based on minute data provided in the Open, High, Low, Close (OHLC) format by Blueshift. It tests the prediction of Bullish Engulfing, Bullish Harami and Bullish Piercing reversal patterns over top Liquid Stocks in the Nifty50 Large Cap and Nifty50 Small Cap. The use of statistical tools including bootstrap sampling, skewness adjusted t-test and the binomial test based on returns and winning probability is taken to accept/reject the hypothesis. The Marshall-Young-Rose (MYR) exit strategy which specifies a prespecified time to exit the market is used by varying the holding time by 1, 5, and 10 minutes.

Introduction

The candlestick patterns are price-action strategies that were developed by Japanese rice traders and became popular in equity markets after being introduced by Nison in 1991. The proposition is to chart the price moments with respect to the open, high, low, and close prices of a trading period. This study focuses on reversal candlestick patterns to predict future turning points on which short-term trades can be made.

There has been extensive work to study the profitability of candlestick patterns across American, Japanese, Taiwanese and Chinese markets since the first empirical study by Caginalp and Laurent in 1998. Despite the popularity, the academic literature gives mixed results regarding the effectiveness and profitability of candlestick patterns. The studies of Caginalp and Laurent (1998), Goo et al. (2007) and Lu and Shiu (2011, 2012, 2014) find some candlestick patterns profitable in short-term investments while on the contrary Lo et al. (2000), Fock et al. (2005), Marshall et al. (2006, 2008) and Horton (2009) failed to suggest any significant returns.

The examination of these studies showcases that the market and stock samples on which the trades are being analyzed might be contributing to the different results arrived at. Hence, it is worth identifying the attributes of the stocks and the market in which these strategies are more suitable.

Liquidity as a parameter is an important factor to consider while selecting the stocks in the Indian Market (NSE) in this study to avoid fluctuations in price due to the quantity of stocks, hence the study focuses on the top 25 liquid stocks in NIFTY50 Large Cap and NIFTY50 Small Cap.

Hypothesis

In order to test the statistical significance of candlestick patterns, the study applied two hypothesis tests for each bullish reversal candlestick pattern,

H_0: $u_i \leq 0$, u : Mean returns of signals in stocks with high liquidity, i ∈ {Bullish Engulfing, Bullish Harami, Bullish Piercing}

A one-tailed bootstrapped skewed t-test is applied and the null hypothesis is rejected if the mean returns are positive implying significant returns and acceptance of alternative hypothesis

H_a: $u_i > 0$, u : Mean returns of signals in stocks with high liquidity, i ∈ {Bullish Engulfing, Bullish Harami, Bullish Piercing}

The other hypothesis test is based on the winning probability of each pattern:

H_0: $p_i \leq 0.5$, p_i : winning probability of signals in stocks with high liquidity, i ∈ {Bullish Engulfing, Bullish Harami, Bullish Piercing}

The binomial test is used in this case to reject the above null hypothesis showcasing better probability than randomness leading to accepting alternative hypothesis:

H_a: $p_i > 0.5$, p_i : winning probability of signals in stocks with high liquidity, i ∈ {Bullish Engulfing, Bullish Harami, Bullish Piercing}

Methodology

Data

Time Period of Data:

The timeframe restriction was over to use data between July 2018 and July 2021. Considering the effect of COVID(suddenly bearish from Feb '20 till April '20 and bullish from May' 20 to July '21) the data of that period was omitted in this study's analysis to avoid unreliability in results.

The data of the 6 months before the testing period was utilized to formulate the groups of stocks used to test the strategy. Initially, the stocks group was identified using the data for the first 6 months i.e. 1st July 2018 to 31st December 2018 and all the candlestick patterns were tested over the first quarter i.e. 1st Jan 2019 and 31st March 2019 to avoid the *look-ahead bias*.

Characteristics of Stocks:

In order to test the hypothesis, the following sets of criteria for selecting the universe of stocks were decided.

1) *Market Capitalization*
   Market Capitalization in conventional literature is described as the total value of all a company's shares of stock.
   a) Large Cap: Large-cap stocks are defined as those having a Market Cap of Rs. 20,000 crores or more
   b) Small Cap: Small-cap stocks are defined as those having below a Market Cap of Rs. 5,000 crores.

2) *Liquidity*
   Liquidity is a necessary aspect while trading in the equity market to avoid the spread of buy-sell especially while trading each minute. As per the findings of Ahn et al (2018), AMIHUD Illiquidity Measure is taken as the method for finding the most Liquid Stocks in a group. AMIHUD captures the lack of liquidity by dividing the daily returns by the daily dollar (Rupee) volume. Essentially, it shows the price change that is triggered by a unit of price volume.

   For the given stock, it is calculated as

$$AMIHUD \;=\; \frac{1}{T}\sum_{t=1}^{T}\frac{|r_t|}{Dollar\,Volume_t}$$

   Where T is the number of days with trading volume and r_t is the return on day t

Stock Groups:

On the basis of these criteria, the universe (groups of stocks) were narrowed down to the following sets.

- **Large Cap, Highly Liquid Stocks: NIFTY50LargeCapLiquid25**
  ○ The large-cap stocks were taken according to Nifty50 index data of the NSE Dec 2018 Report (refer to Appendix)
  ○ The Amihud Measure of Illiquidity was used to identify 25 stocks that were the least illiquid.
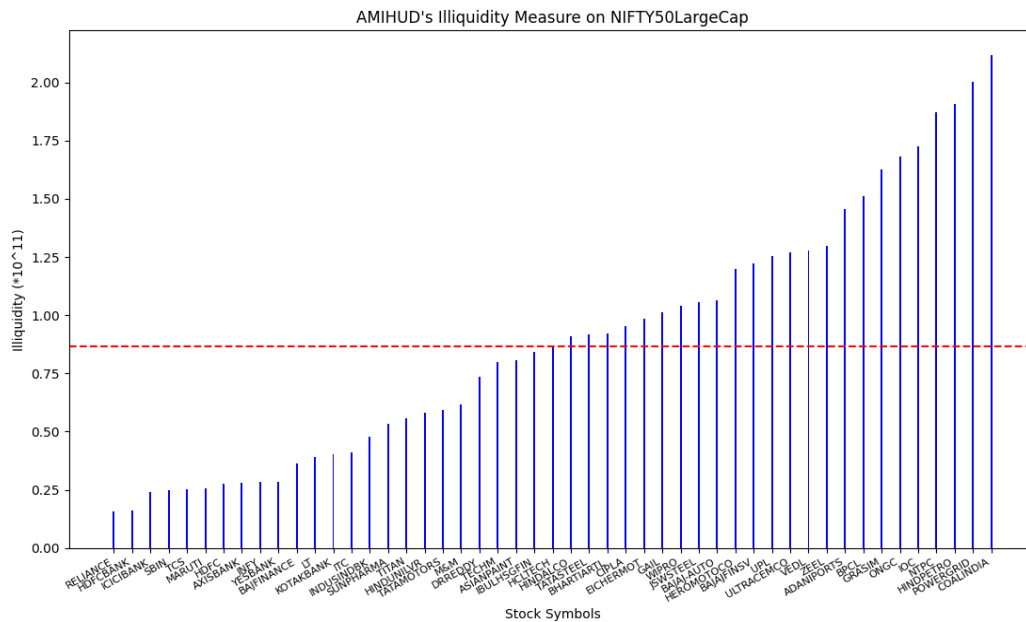
25 Stocks with the lowest illiquidity among Nifty 50 Large Cap were selected

- **Small Cap, Highly Liquid Stocks: NIFTY50SmallCapLiquid25**
  - The small cap stocks were taken according to Nifty50 SmallCap index data of the NSE Dec 2018 Report (refer to Appendix)
  - The Amihud Measure of Illiquidity was used to identify 25 stocks that were the least illiquid.
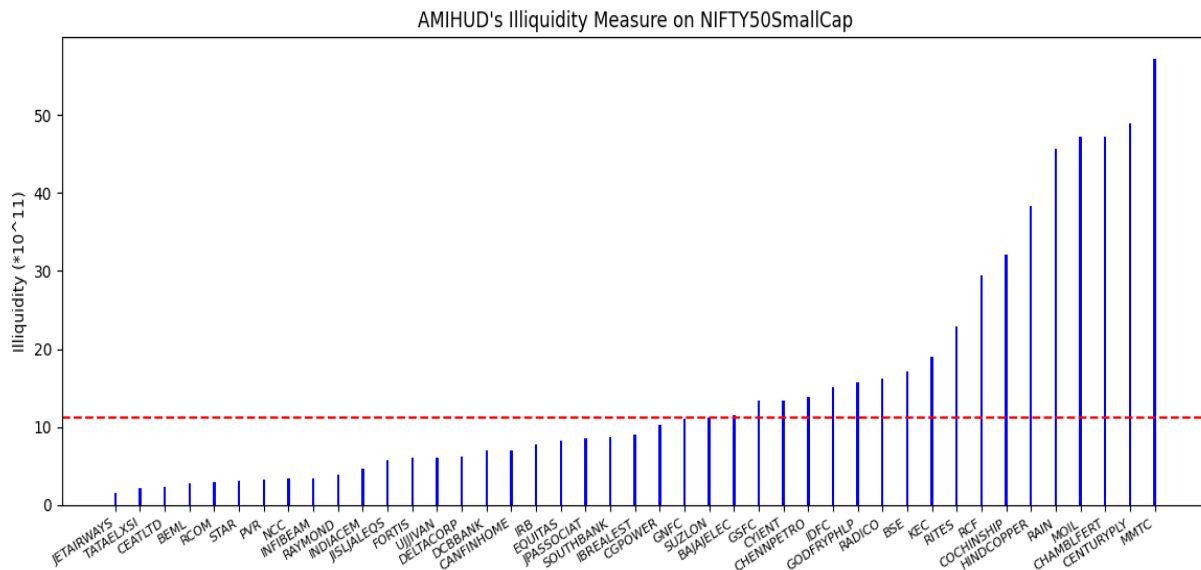
25 Stocks with the lowest illiquidity among Nifty 50 Small Cap were selected

Strategy

Candlestick Pattern:

There are mainly two types of candlestick patterns:
- *Continuation pattern* - It predicts the continuation of an established trend.
- *Reversal pattern* - It indicates a trend change.

Reversal patterns, according to Nison (1991), are more significant since they would enable traders to purchase at the bottom and sell at the top so the study focuses on 3 Bullish CandleStick reversal patterns:

Bullish Engulfing   **Bullish Harami**   **Piercing Line**
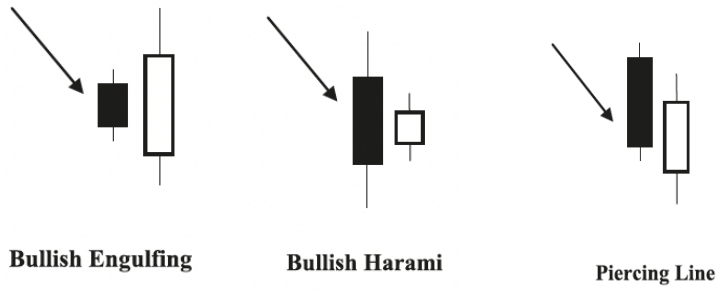
1. Bullish Engulfing: C2 > O1 > C1 > O2 & EMA10 > C2 (Downtrend)
2. Bullish Harami: O1 > C2 > O2 > C1 & RB1 > RB2 & EMA10 > C2 (Downtrend) OR
   C1 > C2 > O2 > O1 & RB1 > RB2 & EMA10 > C2 (Downtrend)
3. Piercing Line: O1 > C2 > L1 > O2 & C2 > 0.5(O1 + C1) & EMA10 > C2 (Downtrend)

Trend Identification:

Since the hypothesis is based on testing candlestick reversal bullish patterns, a downtrend needs to be identified as argued by Morris(2006). The trend EMA10 (Exponential Moving Average over 10 Candles) is employed in the paper as recommended by Marshall et al. (2006) and Marshall, Young, and Cahan (2008). This trend identification is used as the market is in a state of either an uptrend or downtrend, so to determine the downtrend the closing price should be lower than EMA10. The EMA is determined using the formula below, where "C" represents the closing price and N stands for the average period. N is 10 trading candles in this scenario.

$$EMA_{N,t} = (\frac{2}{N+1})(C_t - EMA_{N,t-1}) + EMA_{N,t-1}$$

Evaluation Methods

Return Calculation

In this paper, the strategy aims to buy a stock at its opening price on the candle/minute after a bullish signal is detected. The MYR exit strategy is employed as the stock is held over a fixed period and sold at the closing price. The returns are then calculated for different groups of stocks selected

on the basis of the criteria mentioned above by varying the holding period to 1, 5 and 10 mins. On the basis of this, the optimal holding period is determined for bullish signals. For the MYR exit strategy employed, the holding period continuous return is calculated as -

$$r_{t+h} = lnC_{t+h} - lnO_t$$

"O" and "C" represent the opening price at the time 't' and closing prices at time 't+h' respectively where 'h' is the number of candles in a holding period.

Skewness Adjusted T-test

The profitability of candlestick patterns is tested using the skewness adjusted t-test developed by Johnson(1978). We do not follow the assumption that the data is t-distributed. The data is assumed to have some degree of skewness. The skewness-adjusted t-test is a non-parametric statistical test used to handle biased distributions of return rates. Based on John(1999), the following formula is used to calculate the statistics:

$$t_{sa} = \sqrt{n}(s + \frac{1}{3}\gamma s^2 + \frac{1}{6n}\gamma)$$

$$s = \frac{\bar{r}}{\sigma_r} = standardized\ mean\ of\ returns$$

$$\gamma = \sum_{i=1}^{n} \frac{(r_i - \bar{r})^3}{n\sigma_r^3} = estimated\ skewness\ of\ returns$$

Since the data of individual returns is not following standard t distribution, Confidence intervals of statistical significance 99%,95%, and 90% are calculated using a bootstrapping technique based on Lyon et al (1999).

Bootstrapping Technique: 5000 bootstrapped samples were drawn from the original n/4 samples. For each resampled group, a one-tailed t-test was performed to calculate the t-statistics, t_{stat}. The lower and upper values of the confidence interval are picked up at quantiles of x% and (100-x)% of sorted t_{stat} where x takes up values: 1,5,10.
R statistical software is used to calculate the critical values. The code can be found in the appendix.

Binomial test

In order to investigate the predictive power of candlestick patterns, a binomial test is performed. The number of correct signals divided by the number of total signals gives the winning probability. For candlestick patterns to indicate future short-term returns, it should be greater than 0.5. The interpretation of winning probability varies depending on the exit strategy employed. In our case of the MYR exit strategy, it is the probability that the holding period return would be positive(negative) for a bullish(bearish) signal. In order to test the null hypothesis, the Z statistic is calculated as follows -

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

np: actual number
np_0: expected number of success where $p_0 = 0.5$,
***σ**: standard error.*

The Z statistic has a normal distribution as per the Central Limit Theorem. The standard error is given as follows-

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

The confidence intervals for Z statistics are calculated assuming a normal distribution. The significance values considered are 90%,95%, and 99%.
The Python code used to perform the calculations of returns, T statistic, Winning probability and Binomial Test is given in the appendix.

Risk Management and Assumptions

a. *Commission:*
The commission is not taken into account as the returns calculations are calculated as

$$ln\ (C_{t+h} / O_t)$$

If we assume the number of stocks ordered as q, and assume brokerage as the standard rate of Rs.20, then the return calculation of each will be changed to:

$$ln\ (q. C_{t+h}\ +\ 20/ q. O_t\ -\ 20)$$

The term $q.Q_t$ denotes the transaction value and since
$20 <<< q.O_t$, it can be reframed as:

$$ln\ (C_{t+h}/O_t)\ -\ (20/q. C_{t+h}\ +\ 20/q. O_t\ )$$

Now again since $20 <<< q.O_t$, the second term can be ignored in the sample data and Hence the commission was omitted from the analysis. Also, it is to be noted that, it is only possible since we are trading over highly liquid stocks hence enabling us to trade higher volumes.

b. *Slippage*:
Slippage has been ignored because the stocks chosen are highly liquid stocks, so in normal market conditions it is highly probable that the order gets placed quickly resulting in non-significant slippage which can also be verified by checking the slippage percentage in the transaction files.

c. *Stop Loss:*
The research involves trading stocks at each minute even with holding intervals of a minute. In order to implement the stop loss, the value detected should be checked at each second which was not possible to implement due to the limitation of minimum frequency as 1 minute for callback functions. Therefore, stop loss was omitted as it was not that significant as we traded liquid stocks in small time frames.

Empirical Results

1. **Stock Group:** NIFTY50LargeCapLiquid25

CandleStick	Returns over different Holding Period		
	1 Min	**5 Mins**	**10 Mins**
Engulfing			
n	2277	2276	2276
Returns%	-0.004	-0.057	-0.206
t-stat	-0.086	-0.504	-1.262
Confid. Interval	(-1.668, 1.641)	(-1.778, 1.721)	(-1.707, 1.939)
Winning Prob.	0.485	0.489	0.482
Z	-1.446	-1.048	-1.719
Harami			
n	7626	7620	7616
Returns%	-0.038	0.302	0.215
t-stat	-0.357	1.034	0.535
Confid. Interval	(-1.619, 2.319)	(-1.856, 1.831)	(-1.785, 1.753)
Winning Prob.	0.491	0.510	0.515
Z	-1.512	**1.832****	**2.612*****
Piercing			
n	346	346	346
Returns%	-0.004	0.009	-0.001
t-stat	-0.918	0.807	-0.052
Confid. Interval	(-1.779, 1.769)	(-2.193, 1.706)	(-1.969, 1.864)
Winning Prob.	0.448	0.483	0.477
Z	-1.935	-0.645	-0.860

Note. The "n" represents the number of signals generated by the candlestick for a particular holding period. The "Return%" is the average holding period return. The "t-stat" is obtained by a one-tailed t-test that $\mu \leq 0$ to accept or reject null hypotheses on return on stocks. "Confid. Interval" represents the t_{alpha} to reject the null hypothesis. The "Winning Prob." is the winning probability that the signal generated would earn positive profits. The "Z" is the one-tail Z test used to reject or accept the null hypothesis that winning probability $\leq .5$. *, **, *** denote statistical significance at 10%, 5% and 1%, respectively.

2. **Stock Group:** NIFTY50SmallCapLiquid25

CandleStick	Returns over different Holding Period		
	1 Min	5 Mins	10 Mins
Engulfing			
n	802	802	802
Returns%	0.200	0.074	-0.168
t-stat	**2.384****	0.489	-0.831
Confid. Interval	(-3.054 1.906)	(-2.129, 1.698)	(-1.859, 1.786)
Winning Prob.	0.509	0.510	0.475
Z	0.494	0.565	-1.412
Harami			
n	3740	3738	3737
Returns%	0.018	0.325	0.148
t-stat	0.165	1.304	0.421
Confid. Interval	(-1.964, 1.713)	(-1.968,1.704)	(-1.861, 1.661)
Winning Prob.	0.471	0.491	0.499
Z	-3.597	-1.047	-0.147
Piercing			
n	183	183	183
Returns%	0.043	0.045	0.027
t-stat	**2.931*****	**1.748***	0.742
Confid. Interval	(-5.058, 2.239)	(-2.400, 1.713)	(-1.898 , 1.999)
Winning Prob.	0.557	0.551	0.535
Z	**1.552***	**1.404***	0.960

Note. The "n" represents the number of signals generated by the candlestick for a particular holding period. The "Return%" is the average holding period return. The "t-stat" is obtained by a one-tailed t-test that $\mu \leq 0$ to accept or reject null hypotheses on return on stocks. "Confid. Interval" represents the t_{alpha} to reject the null hypothesis. The "Winning Prob." is the winning probability that the signal generated would earn positive profits. The "Z" is the one-tail Z test used to reject or accept the null hypothesis that winning probability $\leq .5$. *, **, *** denote statistical significance at 10%, 5% and 1%, respectively.

Empirical Analysis

Hypothesis Testing over Returns on Signals
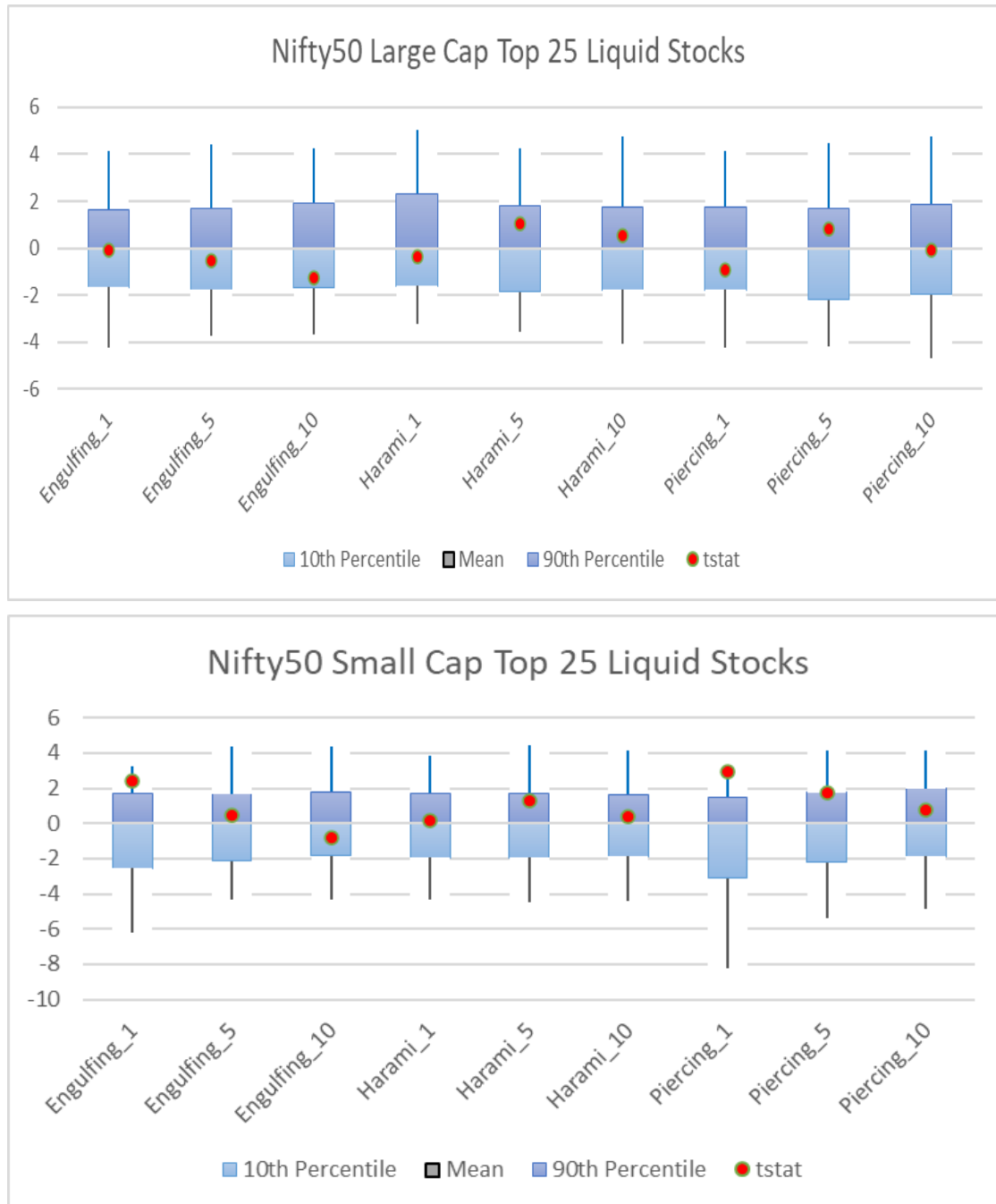
Note. *The image depicts the t-stat and confidence intervals for various candlesticks patterns with different holding periods (Candlestick_Holding period) with the help of a box. The top end of the line represents the maximum (100th percentile), the higher end and lower end of the box represent the confidence interval (90th percentile, and 10th percentile respectively), and the lower end of the line represents the minimum(0th percentile) obtained from bootstrapped sampling. The t-stat is represented by a red dot.*

Hypothesis Testing over Winning Probability

The following graphs depict the Z score obtained for different candlestick patterns with varying holding periods for both the stock groups selected above. The standard values for Z-Score for normal distribution are

- 1.282 : 90% significance level
- 1.645 : 95% significance level
- 2.326 : 99% significance level

Z-score for Top 25 Liquid Stocks in Nifty 50 Large Cap:
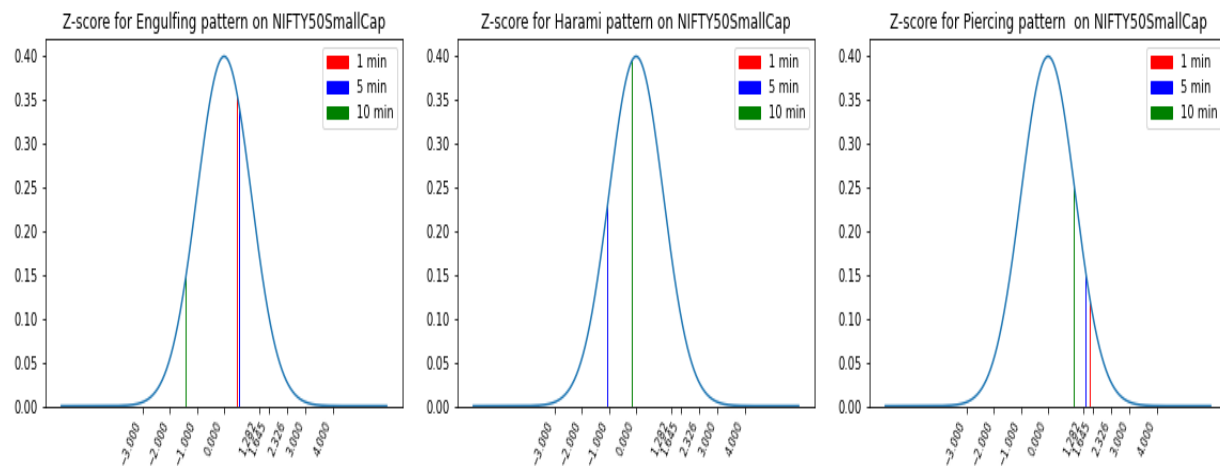
Z-score for Top 25 Liquid Stocks in Nifty 50 Small Cap:
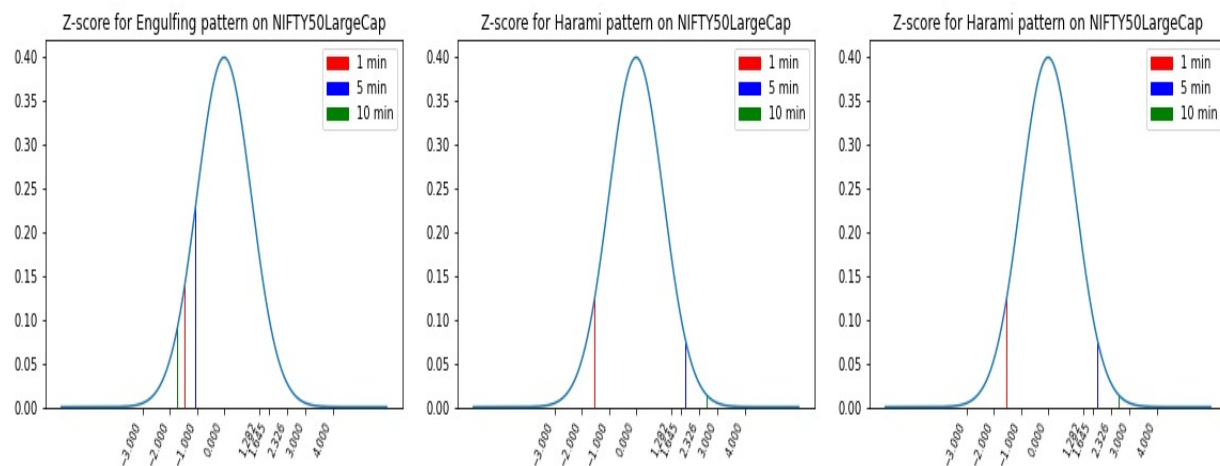
Discussion

This part explains the test statistics and empirical results obtained from the data via. skewness adjusted t-test and binomial test. (Tables 1 & 2 state statistics for two different types of stocks).

In the case of NIFTY50LargeCapLiquid25, none of the patterns and holding period produces any statistically significant returns.

However, in the case of NIFTY50SmallCapLiquid25, the Piercing candlestick pattern with holding period 1, 5 min and Engulfing Pattern with holding period 1-min generate statistically positive returns.
   ● For the Piercing pattern with a holding period of 1 min, the null hypothesis can be rejected with 99% confidence implying that positive returns are generated. The binomial test on winning probability of 0.557(90% confidence) confirms that the majority of trades were profitable and positive returns are distributed among all the profitable trades and not concentrated on some highly profitable trades.
   ● Again in the case of piercing, even when the holding period is 5 min, the null hypothesis can be rejected with 90% confidence and the winning probability is 0.551(90% confidence).
   ● In the case of Engulfing with a holding period of 1 min, t-tests indicate that the null hypothesis can be rejected with 95% confidence but binomial tests on winning probability are not normally significant to reliably predict the market directions.

Conclusion

This study aimed to evaluate the effectiveness of bullish reversal candlestick strategies on liquid stocks of the NSE stock exchange. The idea behind the hypothesis was that highly liquid stocks would show significant results on these price action strategies. The time period of candlesticks was taken as 1 minute. The frequency of trades (i.e. detection of candlestick patterns) across the strategies and stocks for 3 months was in the range of 3 to 60 trades per day, with an anomaly only noticed in the Harami Candlestick pattern in Nifty50LargeCap25.

To effectively test this hypothesis, the universe was then divided on the basis of market capitalization. However, highly liquid large-cap stocks showed no significant results, across all strategies. For highly liquid Small Cap stocks, reversal candlestick patterns except Harami worked with the most significant considering a 1 min holding period.

The highest and most statistically significant returns were generated by the Piercing Candlestick pattern.

Contribution

- *Research on Hypothesis*: Group discussion with whole team
- *Research Methodology of Strategy Hypothesis*: Vineet Ahuja, Pratiksha Jain, Aryan Goyal
- *Collection and Identification of Stocks Universe and data*: Pratiksha Jain, Rohan Bharti
- *Implementing strategy on Blueshift:* Vineet Ahuja, Pratiksha Jain
- *Risk Management:* Vineet Ahuja
- *Research (and implementation) of evaluation criteria of Hypothesis:* Aryan Goyal, Kratika Gupta
- *Empirical Results:* Rohan Bharti, Kratika Gupta
- *Empirical Analysis*: Vineet Ahuja, Pratiksha Jain, Kratika Gupta
- *Discussion and Conclusion*: Group discussion with whole team

Appendix

Code:
- AMIHUD Illiquidity Measure : This Python code implements the AMIHUD illiquidity measure
- BlueShift Strategy Code: This code contains the implementation of the strategy to determines trades and generate transaction data via blueshift
- Evaluation: This python code calculates returns, performs t-test and binomial test on transactions.
- Bootstrapped Skewness-adjusted T-test : This is an R script which runs bootstrapped skewness-adjusted t-test over the samples of returns obtained from the above python code

BlueShift Logs: Available here

Indices Of Data (NSE Dec 2018): They were taken from Nifty Indices. The exact files used in the code are available at this Github Repository

References

[1] Ahn, H.-J., Cai, J., & Yang, C.-W. (2018). Which Liquidity Proxy Measures Liquidity Best in Emerging Markets? *Economies*, *6*(4), 67. https://doi.org/10.3390/economies6040067

[2] Caginalp, Gunduz and Laurent, Henry, The Predictive Power of Price Patterns. Applied Mathematical Finance, Vol. 5, pp. 181-206, 1998, Available at SSRN: https://ssrn.com/abstract=932984

[3] Fock, J.H., C. Klein, and B. Zwergel, 2005, Performance of candlestick analysis on intraday futures data, Journal of Derivatives 13, 28-40.

[4] Goo, Y.-J., Chen, D.-H., & Chang, Y.-W. (2007). The application of Japanese candlestick trading strategies in Taiwan. Investment Management and Financial Innovations, 4(4), 49-79.

[5] Horton, M.J., 2009, Stars, crows, and doji: The use of candlesticks in stock selection, Quarterly Review of Economics and Finance 49, 283-294.

[6] Johnson, N. J. (1978). Modified t tests and confidence intervals for asymmetrical populations. Journal of the American Statistical Association, 73, 536-544.

[7] Lyon, J. D., Barber, B. M., & Tsai, C.-L. (1999). Improved Methods for Tests of Long-Run Abnormal Stock Returns. The Journal of Finance, 54(1), 165–201.

[8] Lu, T.-H. (2014). The profitability of candlestick charting in the Taiwan stock market. Pacific-Basin Finance Journal, 26, 65-78

[9] Lu, T.-H., & Shiu, Y.-M. (2012). Tests for two-day candlestick patterns in the emerging equity market of Taiwan. Emerging Markets Finance & Trade, 48(Suppl. 1), 41-57

[10] Lu, T.-H., Shiu, Y.-M., & Liu, T.-C. (2012). Profitable candlestick trading strategies—The evidence from a new perspective. Review of Financial Economics, 21, 63-68.

[11] Lo, Andrew & Mamaysky, Harry & Wang, Jiang. (2000). Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. Journal of Finance. 55. 1705-1770. 10.1111/0022-1082.00265.

[12] Marshall, B.R., M.R. Young, and L.C. Rose, 2006, Candlestick technical trading strategies: Can they create value for investors? Journal of Banking and Finance 30, 2303-2323

[13] Marshall, B.R., M.R. Young, and R. Cahan, 2008, Are candlestick technical trading strategies profitable in the Japanese equity market? Review of Quantitative Finance and Accounting 31, 191-207.

[14] Morris, G. L. (2006). Candlestick charting explained: Timeless techniques for trading stocks and futures (3rd ed.). New York, NY: McGraw-Hill.

[15] Tharavanij, Piyapas & Siraprapasiri, Vasan & Rajchamaha, Kittichai. (2017). Profitability of Candlestick Charting Patterns in the Stock Exchange of Thailand. SAGE Open. 7. 215824401773679. 10.1177/2158244017736799.

[16] Zhu, M., Atri, S., & Yegen, E. (2016). Are candlestick trading strategies effective in certain stocks with distinct features? Pacific-Basin Finance Journal, 37, 116–127.