# Credibility of Crossovers-An empirical analysis using Logistic Regression

Dr.R. Subathra

Assistant Professor of Statistics, Government Arts College(Autonomous), Salem-7, Tamil Nadu Email – rsubathraarumugam@gmail.com

Abstract: The basic purpose of an average is to smooth the data. In the field of trading, this measure serves several other purposes. The rolling averages serve as a vital source of identifying the support and resistance for prices. The crossovers of the short term and long term rolling averages throw light on the momentum of prices. The Moving Average Convergence and Divergence (MACD) which is an ideal summary of crossovers is commonly used in trading. The rolling averages used to calculate MACD are computed using several approaches. The efficiency of MACD changes with the type of rolling averages are computed using the Simple Moving Average (SMA) and the Exponential Moving Average (EMA). An ideal approach is the one which generates more returns. Since it is proved that any technical indicator generates false signals, the credibility of the two strategies is tested with Logistic Regression. While doing so, the risk involved in a strategy is found using accuracy of the predictive models. The study identifies that EMA has more credibility in generating better returns with more accuracy.

**Key Words:** Average True Range, Crossovers, Confusion Matrix, Exponential Moving Average, Logistic Regression, Moving Average Convergence and Divergence, Null Deviance, Relative Strength Index, Simple Moving Average.

## **1. INTRODUCTION:**

In this modern world of communication revolution, trading in stock market is a resourceful experience mainly due to availability of all price movements at the trader's desk. The same is a reason for worry because the data is voluminous and the trader requires many tactics to predict the movement. Among a bunch of tactics in technical analysis, the most commonly used is the average of data. The process of averaging which plays a vital role in smoothing the data has evolved as many tools like rolling averages, crossovers and moving average convergence and divergence (MACD). The averages when used in the form of MACD add more colours to the data. A positive MACD indicates upward momentum while the negative MACD indicates the downward momentum. Such a resourceful tool has many issues with it. The first one being the time frame to be used to compute the rolling averages. There is no common threshold and the best time frame should be identified based on experience. The second concern is regarding the type of rolling average to be used in the process of computing MACD. While there are different methods of computing the rolling averages, this study attempts to compare the potential of MACD with respect to Simple Moving Average (SMA) and Exponential Moving Average (EMA). According to the experience, success in trading depends on the usage of an optimum combination of technical tools. Hence the two alternative MACDs used in this work are analyzed in combination with the Relative Strength Index (RSI) and the Average True Range (ATR). Even though the technical analysis claims that the share price movements are systematic, many empirical studies have proved the presence of randomness in their behaviour. It is a well known fact that the technical analysis generates false signals. Hence the credibility of any strategy should be checked with a statistical model. Many classification methods like Bayes classifier, K-nearest neighbour, Decision trees, discrimnant function and Logistic Regression exist in the literature. Among these methods the Binary Logistic regression is identified as the most suitable method for the present study. The best strategy is identified as the one which generates more returns. The risk associated with the returns is computed using the accuracy of the Logistic Regression model framed for the two different strategies. Before computing the accuracy of the signals provided using Logistic Regression, the credibility of the Logistic Regression model is tested by comparing it with the null model. Hence after many such tests, the study attempts to find the better strategy for trading.

## 2. LITERATURE REVIEW:

Identifying the potential trading points demands greater expertise and innovative application of technical analysis tools. The moving average and its various enhanced forms are generally used by the traders. The following studies on two different strategies based on rolling averages form the foundation for this work.

Pring (1991) [1] argued that Technical analysis is actually a manifestation of the idea that prices follow trend. He advocates a combination of technical tools as no single indicator has the ability identify the trend reversal.

Pruitt et al (1992)[2] and Pruitt & White (1988) [3] evaluated the profitability using a combination of Relative Strength Index, Moving averages and Cumulative volume.

Murphy (1999) [4] states that the Technical analysis is a blend of many approaches and the trader can beat the market using the clues given by the technical tools. He proved that the more the trader uses indicators, the more he may be able to choose the better clues and consequently has more chances to earn abnormal returns.

Rodriguez et al. (2000) [5] elucidated the profit generation ability of simple technical trading strategy employing the Artificial Neural Network (ANN). In the absence of transactional cost, the strategy based on Technical analysis produces greater return in contrast to B&H policy. He proved that the trading rule loses its ability of abnormal profit generation when the market is "bullish" and thus traders with buy and hold strategy receive greater returns.

Ellis, C. A. and S. A. Parbery (2005) [6] found that the about three out of four cases trading rule signals are false giving an implication that leaves a lot of space to improve the trading rule performance.

Ming-Ming, L. and L. Siok-Hwa (2006) [7] established that the Adaptive Moving Average over fixed length responds to the price movements.

Chang et al. (2006) [8] employed the moving average approach in Taiwan stock market and observed the excess profit as compared to the buy-and-hold strategy even after considering the transaction cost.

Vasiliou et al. (2006) [9] conducted the study by using Moving average and moving average convergence divergence (MACD) rules and concluded that these strategies produced above average returns as compared to B&H strategy.

To investigate the question that whether the tools of Technical analysis outperformed the B&H policy, Lento and Gradojevic (2007) [10] conducted a study employing MACD, Bolinger Bands and filter rules on four different indexes. They established that out of the four rules, the filter and MACD rules performed well. Similarly BB and filter rules are not profitable after considering the cost of transactions.

Metghalchi, M., X. Garza-Gomez, et al. (2008) [11] have proved that the profitability with Variable Moving Averages and Fixed Moving Averages in China, Thailand, Taiwan, Malaysia, Singapore, Hong Kong, Korea, and Indonesia stock markets.

Milionis, A. E. and E. Papanagiotou (2008) [12] proved that the performance of the moving average trading rule is improved, if it is combined with other indicators.

The study conducted by Khan et al. (2016) [13] investigated the predictability of moving averages individually as well as with the combination of relative strength index (RSI) and stochastic RSI on Karachi Stock Exchange data. The study established that the predictability of moving averages increases in the presence of these oscillators. The use of technical analysis outperformed the buy and hold strategy in generating abnormal returns.

Using logistic Regression model Subathra.R(2020) [14] concluded that the usage of DeMark's pivot point is mor e stable than pivots computed using Standard and Woodie's approach. In this work the ideal pivot is identified by usin g three types of pivots in conjunction with ATR, RSI and MACD.

## 3. RESEARCH METHODOLOGY:

Moving averages smooth the price data by filtering the noise to form a trend following indicator. A simple moving average is formed by computing the average price of a security over a specific number of periods. In general they are computed based on closing prices. It is called moving or rolling because the old data is dropped as new data becomes available, causing the average to move along the time scale. It is a lagging indicator as it is based on past prices.

Exponential moving averages (EMAs) reduce the lag by applying more weight to recent prices. The EMA for the current period depends on the EMA of the previous period. To start with the SMA is taken as the initial EMA value. After calculating the weight multiplier, the current EMA is computed using the current price, previous EMA and the weight multiplier.

The main aim of this work is to identify an appropriate trading strategy. But using a single indicator as a market monitor may not be an effective practice. Hence multiple indicators are used in this work to identify the more competing strategy. A multi-indicator strategy may become redundant when they provide same type of information. Selection of one indicator from each broad category of technical indicators may be an effective way of avoiding this fallacy. With this realization this work considers the following technical indicators to frame the strategies.

The MACD (Moving Average Convergence/Divergence) is a technical analysis tool which shows the relationship between prices and rolling averages. It is the difference between 26 period and 12 period rolling averages. The 9 period rolling averages indicate the trading positions. The divergence in MACD indicates the altering trend. Although the standard setting for the MACD considers the Exponential moving average, any type of moving average can be used. The types used in general are: The Simple Moving Average(SMA), the Exponential Moving Average (EMA), the Weighted Moving Average(WMA) and the Adaptive Moving Average(AMA). In this study SMA and EMA are compared.

Relative Strength Index (RSI) is a leading indicator that measures the speed and the price movement. RSI oscillates between zero and 100. RSI is a kind of momentum oscillator which is a number between 0 and 100. The value of RSI is considered overbought when above 70 and oversold when below 30.

The Average True Range is considered as an accurate volatility measure. It measures the intensity of movement of an asset in the past.

With the above technical indicators, the study considers two strategies based on SMA and EMA which are:

- 1. MACD using SMA in conjunction with RSI and ATR (MACD-S)
- 2. MACD using EMA in conjunction with RSI and ATR (MACD-E)

The technical indicators claim that specific patterns will lead to movement of stock prices in specific directions. But the literature has proved that all those indicators have uncertainty to certain extend. It is this reason which motivates many researchers to use probabilistic models to make decisions based on technical indicators. The technical analysis tools have enhanced the usage of probabilistic models by throwing more light on the movement of prices. This can be successfully facilitated by the classification models. Many classification methods like Bayes classifier, K-nearest neighbour, Decision trees, discrimnant function and Logistic Regression exist in the literature. Among these methods the Binary Logistic regression is the most suitable method for the present study. The binary Logistic regression model is an enhancement of linear regression. The logistic model best suits the situation in which the dependent variable is dichotomous. In the present study, the dichotomous dependent variable assumes two values: Buy coded as 1 and Sell which is coded as 0. Further the Logistic regression model does not necessitate that the relationship should be linear and the variables should be normally distributed.

The Binary logistic regression performs the same task of Linear Discrimnant Analysis but it uses a Sigmoid function that provides an output between 0 and 1. This aspect makes it appropriate for financial studies on stock market movements and Bankruptcy. The Logistic model uses a probabilistic method based on maximum likelihood estimators with no parametric assumptions. In this point of view, the Logistic regression is more robust method. The model for Logistic regression is

For two classes of output Y, the parameters  $\beta_0, \beta_1, ---\beta_p$  are estimated using Maximum Likelihood estimation. The Logit is given by

$$G(x) = \log \frac{\pi(x)}{1 - \pi(x)} = \log \frac{P(Y=1 \text{ given } X=x)}{P(Y=0 \text{ given } X=0)} = \beta_0 + \sum_{i=1}^p \beta_i X_i - \dots - \dots - (2)$$

The curve of  $\pi(x)$  is called Sigmoid. It is because it results in a S-Shaped nonlinear curve. Thus the model introduces an appropriate link function in the analysis. This model is more relevant when the dataset is very large. This model predicts the logit of Y from X. The logit is the natural logarithm of the odds ratio. The odds ratio is given by

The credibility of the fitted Logistic model is tested using various approaches. In this study the following tests are used to analyze the suitability of the fitted models.

- In the linear regression model, the coefficient of determination summarizes the proportion of variance in the dependent variable explained by the explanatory variables. For regression models with categorical dependent variables, instead of the coefficient of determination, three methods namely Cox and Snell's method, Nagelkerke's method and McFadden's method are used. This study uses McFadden's approach to analyze the suitability of the model. In this approach, -2LL is used to assess the overall fit of the model. The value of this Pseudo R-Square lies between 0 and 1. The value 1 implies a perfect model and the value in the range 0.2-0.4 is considered good.
- In this study, the proposed model is compared with a null model which is an intercept-only model. By deviance we mean the level to which the likelihood of the model exceeds the likelihood of the other model. If the deviance of the null model is low, the interpretation is that an intercept alone is sufficient to create a model. On the other hand, if the residual deviance is low, the implication is that the proposed model is more appropriate. To do this the Likelihood Ratio Test (LRT) is used. LRT is a test of the sufficiency of a smaller model versus a more complex model. The null hypothesis of the test states that the smaller model provides as good a fit for the data as the larger model. If the null hypothesis is rejected, then the alternative, larger model provides a significant improvement over the smaller model.
- Evaluating the accuracy of the fitted models is the main part in the statistical analysis. It is a usual method to test the performance of the model in the past. It is done on the assumption that a model which performs well in the past will also perform well in the future. The classification table gives more details on the

performance of the model. It displays four types of combinations of actual and forecasted values: Actual decreasing and Forecasted decreasing (Correct forecast), Actual decreasing and forecasted increasing (Forecast error of Type I), Actual increasing and Forecast decreasing (Forecast error of Type II), Actual increasing and Forecast). With this confusion matrix, the accuracy of the fitted model is evaluated.

## 4. ANALYSIS:

The study uses the Logistic Regression technique to predict stock price movement. The daily prices of randomly selected NIFTY stocks from 01-01-2020 to 30-04-2021 collected from the official website of National Stock Exchange are used in this study. 75% of the observations are used as training data and the remaining 25% as the testing data for Logistic Regression.

The stocks selected at random from the NIFTY stocks are: PETRONET, HINDPETRO, HDFCBANK, GRASIM and TCS. The analysis is carried out with the following steps.

Step-1: At the outset, the returns generated are computed using three strategies:

- 1. Buy and hold strategy
- 2. MACD-S which is the combination of MACD computed using Simple Moving Average used along with RSI and ATR.
- 3. MACD-E which is the combination of MACD computed with Exponential moving averages, RSI and ATR.

The comparative movement of buy and hold return and the respective strategic return are represented in Figure 1 to Figure 10. The returns are tabulated to identify the prospective strategy.

Step-2: Having computed the returns, the next step is to apply Logistic Regression, the purpose for which is two-fold:

- i. To assess the efficiency of the strategies.
- ii. To compute the risk associated with the returns in terms of the accuracy of the fitted models.

The credibility of the fitted model is tested using the following criteria:

- Likelihood Ratio Test (LRT) is used for comparing the fitted model with the null model. If the proposed model has a deviance less than the null deviance, then the proposed model is the best.
- If the McFadden's R-squared value lies between 0.2 and 0.4, the fitted model is good.
- Back testing the model is done by using 75% of the data as training data and the remaining 25% as the testing data.

Among the two strategies MACD-S and MACD-E, the better strategy is the one which satisfies the above criteria. The Figure-1 and Figure-2 compare the Buy and Hold return with the returns due to the strategies MACD-S and MACD-E for PETRONET.

## Figure-1: RETURNS DUE TO BUY\_HOLD Vs MACD-S for PETRONET



Cumulative percentage of return from Buy\_hold : -6.64 % Cumulative percentage of return from MACD-E Strategy: -2.0 %



Figure-2: RETURNS DUE TO BUY\_HOLD Vs MACD-E for PETRONET

Cumulative percentage of return from Buy\_hold : -6.64 % Cumulative percentage of return from MACD-E Strategy: 7.76 %

The Figure-3 and Figure-4 compare the Buy and Hold return with the returns due to the strategies MACD-S and MACD-E for HINDPETRO.



Figure-3: RETURNS DUE TO BUY\_HOLD Vs MACD-S for HINDPETRO

Cumulative percentage of return from Buy\_hold : -9.52 % Cumulative percentage of return from MACD-S Strategy: 5.75 %



Figure-4: RETURNS DUE TO BUY\_HOLD Vs MACD-E for HINDPETRO

Cumulative percentage of return from Buy\_hold : -9.52 % Cumulative percentage of return from MACD-E Strategy: 20.29 %

The Figure-5 and Figure-6 compare the Buy and Hold return with the returns due to the strategies MACD-S and MACD-E for HDFCBANK.

Figure-5: RETURNS DUE TO BUY\_HOLD Vs MACD-S for HDFCBANK



Cumulative percentage of return from Buy\_hold : 15.17 % Cumulative percentage of return from MACD-S Strategy: 7.86 %



Figure-6: RETURNS DUE TO BUY\_HOLD Vs MACD-E for HDFCBANK

Cumulative percentage of return from Buy\_hold : 15.17 % Cumulative percentage of return from MACD-E Strategy: 28.6 %

The Figure-7 and Figure-8 compare the Buy and Hold return with the returns due to the strategies MACD-S and MACD-E for GRASIM.

## Figure-7: RETURNS DUE TO BUY\_HOLD Vs MACD-S for GRASIM



Cumulative percentage of return from Buy\_hold : 83.0 % Cumulative percentage of return from MACD-S Strategy: 93.31 %



Figure-8: RETURNS DUE TO BUY\_HOLD Vs MACD-E for GRASIM

Cumulative percentage of return from Buy\_hold : 83.0 % Cumulative percentage of return from MACD-E Strategy: 97.76 %

The Figure-9 and Figure-10 compare the Buy and Hold return with the returns due to the strategies MACD-S and MACD-E for TCS.



## Figure-9: RETURNS DUE TO BUY\_HOLD Vs MACD-S for TCS

Cumulative percentage of return from Buy\_hold : 46.75 % Cumulative percentage of return from MACD-S Strategy: 62.39 %



Figure-10: RETURNS DUE TO BUY\_HOLD Vs MACD-E for TCS

Cumulative percentage of return from Buy\_hold : 46.75 % Cumulative percentage of return from MACD-E Strategy: 60.99 %

The results generated in the above analysis are summarized in Table-1

Stock	<b>Buy and Hold</b>	MACD-S	MACD-E
PETRONET	-6.64%	-2%	7.76%
HINDPETRO	-9.52%	5.75%	20.29%
HDFCBANK	15.17%	7.86%	28.6%
GRASIM	83%	93.91%	97.76%
TCS	46.75%	62.39%	60.99%

### Table-1: Cumulative percentage of returns due to Buy and Hold, MACD-S and MACD-E

According to Table-1, the cumulative percentage of return due to MACD-E is higher than the respective returns due to buy and hold strategy and MACD-S except for TCS. For TCS the cumulative percentage of return is higher f or MACD-S

The following table gives the deviance values for the Logistic Regression models which are fitted with the tr ading signals (1 for Buy and 0 for Sell) as the dependent variables and RSI and ATR as explanatory variables.

Stock	Strategy	Null Deviance	<b>Residual Deviance</b>
Petronet	MACD-S	305.61	220.79
	MACD-E	303	301
HindPetro	MACD-S	294.49	222.97
	MACD-E	296.30	208.65
HDFC Bank	MACD-S	306.37	259.98
	MACD-E	337.83	223.86
GRASIM	MACD-S	291.50	252.90
	MACD-E	295.42	216.40
TCS	MACD-S	286.97	267.94
	MACD-E	274.42	232.98

Table-2: Results of LRT of the proposed model with null model

For all the models the null deviance is greater than the deviance of the proposed model and hence testing the efficiency of the strategies using the Logistic Regression model is correct.

After justifying the usage of Logistic Regression model, the efficiency of the model is identified using McFadden's Pseudo R-Squared value. The table-3 gives the summary of the R-Squared values for the fitted models.

Stock	Pseudo R-Squared MACD-S	Pseudo R-Squared MACD-E	
Petronet	0.206768	0.277531	
HindPetro	0.242869	0.295811	
HDFC Bank	0.151400	0.337358	
GRASIM	0.132435	0.267462	
TCS	0.166318	0.251003	

### Table-3: MCFadden's R-Squared value:

The R-Squared values in Table-3 lie between 0.2 and 0.4 for most of the models. For HDFCBANK, GRASIM and TCS, the R-Squared values for MACD-S imply that the models are not good. With the understanding that the R-Squared values for all MACD-E models lie between 0.2 and 0.4, MACD-E is considered a better strategy. The credibility of MACD-E is further tested using back testing for which the trained model fitted with 75% of the data is tested with the remaining 25% of the data. The results are summarized in Table-4.

Stock	MACD-S		MACD-E	
	Confusion Matrix	Accuracy	Confusion Matrix	Accuracy
Petronet	$\begin{pmatrix} 30 & 16 \\ 9 & 28 \end{pmatrix}$	0.70	$\begin{pmatrix} 33 & 11 \\ 10 & 29 \end{pmatrix}$	0.75
HindPetro	$\begin{pmatrix} 33 & 13 \\ 11 & 26 \end{pmatrix}$	0.71	$\begin{pmatrix} 29 & 14 \\ 6 & 34 \end{pmatrix}$	0.76
HDFC Bank	$\begin{pmatrix} 30 & 16 \\ 9 & 28 \end{pmatrix}$	0.70	$\begin{pmatrix} 31 & 8 \\ 11 & 23 \end{pmatrix}$	0.77
GRASIM	$\begin{pmatrix} 22 & 17 \\ 9 & 35 \end{pmatrix}$	0.69	$\begin{pmatrix} 27 & 13 \\ 6 & 37 \end{pmatrix}$	0.77
TCS	$\begin{pmatrix} 14 & 31 \\ 0 & 38 \end{pmatrix}$	0.63	$\begin{pmatrix} 15 & 22 \\ 0 & 46 \end{pmatrix}$	0.73

#### **Table-4: Confusion Matrix and Accuracy**

## 5. FINDINGS:

For all the stocks the accuracy is greater for MACD-E. Even though the return due to MACD-S for TCS is slightly greater, the accuracy of the model is very less compared to the respective MACD-E. Hence based on the returns generated( as summarized in Table-1), deviance of the models( as summarized in Table-2), R-Squared values( as summarized in Table-3) and the Accuracy of the models( as summarized in Table-4), the study concludes that MACD-E is better than MACD-S.

## 6. CONCLUSION :

The study considers two strategies MACD-S and MACD-E which are used in conjunction with RSI and ATR. The two strategies are tested based on the following:

- The cumulative percentage of returns as compared with Buy and Hold strategy. The MACD-E resulted in greater returns than Buy and Hold strategy and MACD-S for all the stocks considered in this study except TCS.
- The credibility of the generated trading signals are tested using Logistic Regression models fitted with trading signals as dependent variable and RSI and ATR as explanatory variables. The suitability of the fitted models is tested with the deviance values. The study highlighted that MACD-E produced minimum deviances.

- The suitability of the models is analyzed using McFadden's R-Squared values which lie in the required range for MACD-E models for all the stocks.
- The accuracy of the model derived using the confusion matrix is greater for MACD-E for all the stocks.

Hence based on the above findings it is concluded that MACD-E generates more returns than MACD-S. The confusion matrices reveal that the false trading signals generated by MACD-E are comparatively less than that of MACD-S. The accuracy of the fitted model is considered as a proxy for the risk of a trading strategy. Hence MACD-E generates more returns with less risk.

## **REFERENCES:**

- 1. Pring, M. J. (1991). Technical analysis: Explained. new york: McGraw-Hill Co.
- 2. Pruitt, S., Tse, K., & White, H. (1992). The CRISMA Trading System: The Next Five Years. *Journal of Portfolio Management*, 18, 22-25.
- 3. Pruitt, S., & White, R. (1988). The CRISMA Trading System: Who Says Technical Analysis Can't Beat the Market? *Journal of Portfolio Management Spring*, 55-58.
- 4. Murphy, J. J. (1999). *Technical Analysis of the Financial Market: A Comprehensive Guide to Trading Methods and Applications*. New York: New York Institute of Finance
- 5. Rodriguez, F., Martel, C., & Rivero, S. (2000). On the profitability of technical trading rules based on artificial neural networks:: Evidence from the Madrid stock market. *Economics Letters*, 69(1), 89-94.
- 6. Ellis, C. A. and S. A. Parbery (2005). "Is smarter better? A comparison of adaptive, and simple moving average trading strategies." *Research in International Business and Finance* 19(3): pp.399-411.
- 7. Ming-Ming, L. and L. Siok-Hwa (2006). "The profitability of the simple moving averages and trading range breakout in the Asian stock markets." *Journal of Asian Economics* 17(1): pp.144-170.
- 8. Chang, Y., Metghalchi, M., & Chan, C. (2006). Technical Trading Strategies and Cross-national Information Linkage: The case of Taiwan Stock Market. *Applied Financial Economics*, *16*, 731-743.
- 9. Vasiliou, D., Eriotis, N., & Papathanasiou, S. (2006). How Rewarding Is Technical Analysis? Evidence from Athens Stock Exchange. *Operational Research*, *6*, 85-102..
- 10. Lento, C., & Gradojevic, N. (2007). The Profitability Of Technical Trading Rules: A Combined Signal Approach. *Journal of Applied Business Research*, 23(1), 13-28.
- 11. Metghalchi, M., X. Garza-Gomez, et al. (2008). "Are moving average trading rules profitable? Evidence from the Mexican stock market." *Journal of Applied Business Research* 24(1): pp.115-128.
- 12. Milionis, A. E. and E. Papanagiotou (2008). "On the use of the moving average trading rule to test for weak form efficiency in capital markets." *Economic Notes* 37(2): pp.181-201.
- 13. Khan, M. A., Khan, N., Hussain, J., & Shah, N. H. (2016). Technical Analysis: Concept or Reality? *Pakistan Business Review*, 18(3), 732-751.
- 14. Subathra.R (2020). A Comparative study on the methods of computing the pivots using Logistic Regression. International Journal on Recent Trends in Business and Tourism (IJRTBT), 4(4),9-15.