Predicting Stock Returns of Infosys

Vaishali Agrawal*
***Research Scholar, IIS (deemed to be University), Jaipur**

Abstract: Forecasting of stock returns is and will always be a vitally important financial notion confronted by investors. Their exist fluctuations in stock returns and investors are always keen to show their interest as they want to take the advantage of potential returns from the organization by way of investing in stocks. Hence, it becomes a matter of concern for investors to predict future stock returns so that they can attain their objective of wealth maximization. This reason creates an urge to explore forecasting of stock returns empirically. This research paper employed ARIMA methodology, developed by Box and Jenkins in 1970, which rely on the previous values of the variable itself. In the paper, this methodology is applied on the stock returns of one of the top IT companies listed on NSE i.e. Infosys Ltd. Data of daily return was collected from 1 April 2008 till 31st March 2018. Results concluded that ARIMA model had strong capability of forecasting in short run.

Keywords: ARIMA, Stock Returns, Forecasting.

INTRODUCTION

Forecasting stock returns is always been a topic for discussion in contemporary financial literature. Investors try hard to contemplate possible future returns of a company's given common stock. The question arises how to identify a probable closest return of a given stock. Traditionally, researchers have made an effort to forecast the stock returns by studying the factors which affects given firm's value or profitability. In the present research paper, an effort has been made to envisage our variable by way of the lagged values of the variable itself. Based on the popular notion of letting the data speak for itself (Gould,1981). Therefore, ARIMA (Auto Regressive Integrated Moving Average) technique has been applied to forecast the stock returns on the basis of its previous values and error term.

There have been many studies conducted on different sectors that have applied ARIMA model for prediction of various time series variables which might include stock prices as well. However, fewer studies have been conducted on IT sector to envisage stock returns of IT company using ARIMA model. More specifically, no study, as per review of literature, has been done using the daily stock returns of IT Company. The present work initiates to fill this gap by taking daily stock returns of one IT sector company in India i.e. Infosys.

Infosys Ltd was incorporated in 1981 as Infosys Consultants Pvt Ltd by Mr. N.R Narayana Murthy in Bangalore, Karnataka. It is India's second largest and world's 596th largest public IT Company based on revenue. Initially, seven engineers established Infosys in Pune, Maharashtra with an investment of $250 in 1981. In April 1992, company changed its name to Infosys Technologies Private Limited. Thereafter, again in June 1992 after becoming a public limited company named itself as Infosys Technologies Limited. Lastly, in June 2012 it renamed as Infosys Limited (Information Systems). In the beginning year's company's revenue reached to US$ 100 million which increased to US$ 1 billion in 2004 and US$ 10 billion in 2017-18. The Company was the first Indian company to be listed on the NASDAQ in the year 1999. On Bombay stock exchange its shares were listed for the first time in February 1993 with trading opening at 145 per share.

REVIEW OF LITERATURE

Afeef, M & et.al (2018) employed ARIMA methodology to forecast stock prices of a Pakistan based company namely Oil & Gas Development Company Limited (OGDCL). The researcher considered daily adjusted closing stock prices of OGDCL for almost 15 years starting from 2004 till 2018 with 3632

observations. Results depicted that for the purpose of prediction in short-run, ARIMA modeling has great potential. Consequently, it was suggested for investors to consider the findings of the study to supplement their aptitude of forecasting. **Gay (2016)** made an effort to investigate the relationship of macroeconomic variables on stock returns of BRIC countries that include Brazil, Russia, India and China. He made use of the Box-Jenkins method to serve the purpose. The factors taken into account were the exchange rates and the oil prices. No statistically significant association was found to be there between the given macroeconomic factors and stock returns for any of the BRIC economies. Moreover, no significant link was identified of stock return with its lagged values for any of the four countries. **Gupta, S & Kashyap, S. (2015)** did endeavour to generate prediction of exchange of Indian currency vis-à-vis USD, GBP, YEN and EURO. They applied Box-Jenkins methodology (ARIMA) on the collected data of twelve months starting from April 2014 to March 2015. In addition, they applied ADF test to verify the stationarity of data and the results were attained at first difference in the data. The best terms of AR and MA were selected after exercising various models and finally best-fit model was selected on the basis of minimum SBC, AIC, sum of squared errors and Q statistics to predict the respective currencies. The forecasted results can provide insights to policymakers, forex dealers, government, corporate etc. to design policies along with generation of prediction in the desired time period. **Hamjah (2014)** also used ARIMA for prediction of rice production in Bangladesh. He compared the actual data of rice production with the predicted values and concluded that model had a very short run prediction capability. **Mondal, P & et.al. (2014)** studied 56 stocks from seven different sectors listed on NSE. Researchers collected data of 23 months for the empirical study. Further, AIC was used to select the best ARIMA model. Results indicated that ARIMA provides best accurate results as above 85% of predictions using ARIMA model for all sectors were accurate. Moving to specific sectors, forecasting of FMCG sector was more accurate as compared to the predictions for Banking and Automobile sectors. **Devi, B & et.al. (2013)** selected top four companies out of which Nifty Midcap 50 was selected on the basis of having maximum midcap value for analysis. The historical data of selected companies for past five years was collected and trained by applying ARIMA model with different parameters. Further, the accuracy of predicted results was checked using criteria like AIC and BIC. Lastly, analysis of trained model was conducted to find the market behaviour and trend for future forecast.

OBJECTIVES

- To forecast the stock returns of Infosys company.
- To analyze the variation in actual and forecasted stock returns of Infosys.
- To check the applicability of ARIMA model in predicting stock returns of Infosys.

ARIMA MODEL

ARIMA model is explained in Box-Jenkins methodology. This methodology is used to identify a potential model out of general class of models. In addition, it uses both the techniques of autoregressive (AR) and moving average (MR) for forecasting as well as tries to search best combination of two techniques. This model was initially introduced by two statisticians namely George P Box and Gwilym Jenkins in their book "Time Series Analysis: Forecasting and Control" (Box & Jenkins, 1970).That's why it is also known as Box-Jenkins methodology. In order to get better results from the model, it is suggested by researchers to work on at least 100 observations or more.

ARIMA models are generally expressed like "ARIMA(p,d,q)" , here the three terms are defined as follows:

- "AR" in ARIMA is called **Autoregressive term** in the model.
- "I" in ARIMA is called the **Integrated** feature of a time series."I" also takes care of differencing to make a time series stationary.

- "MA" in ARIMA represents **Moving Average** term in the model. It assumes that a time series is a function of its errors.

The creators of the model, Box and Jenkins, have focused on the principle of parsimony which emphasize on keeping the model as simple and concise as possible. For the prediction of time series, both the model developers proposed a four tier model. The four steps of ARIMA model are:

1) Model Identification
2) Parameter estimation
3) Diagnostic Checking
4) Forecasting

RESEARCH METHODOLOGY

While using time series econometric framework, it is advisable to extricate information related to a variable which can be gathered from the variable itself (Asteriou & Hall, 2007). The general equation of an ARMA model (Asteriou & Hall, 2007) is as follows:

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \text{-} + \varphi p Y_{t-p} + \varepsilon t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \text{-} + \theta q \varepsilon_t\text{-}q$$

Here, Y_t is the predicted value of the variable, $Y_{t-1,}$ Y_{t-2},…..Y_{t-p} are the lagged values of the autoregressive term(AR), ε_t is the error term, ε_{t-1}, ε_{t-2}, - - - , ε_{t-q} are the lagged values of the moving average(MA) or error terms, φ and θ are the coefficients of the regressors.

Applying ARMA process on non stationary data will definitely provide no results. Therefore, the most appropriate and efficient ARIMA model was applied on daily stock returns of TCS collected from 1 April 2008 till 31st March 2018 in order to forecast more accurate results of stock returns.

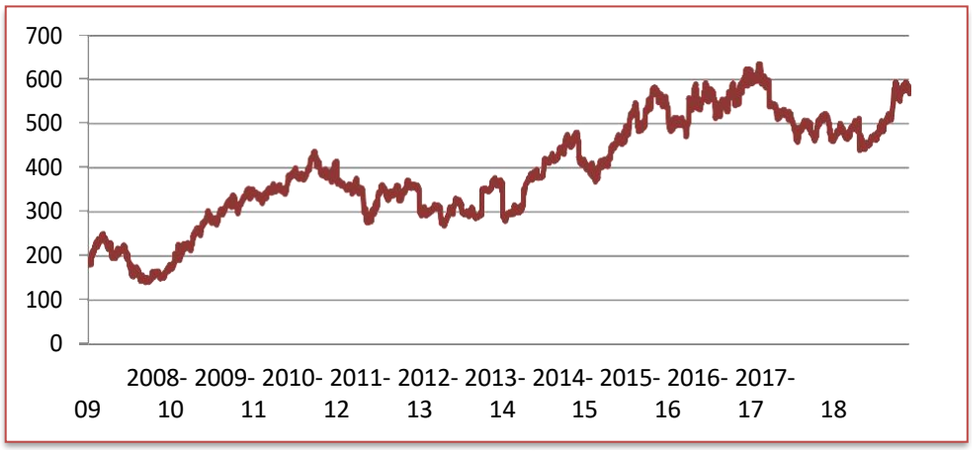
DESCRIPTIVE STATISTICS

	INFOSYS
Mean	0.0467
Maximum	15.5149
Minimum	-23.9012
Std. Dev.	1.9107

Table-1: Descriptive Statistics

Further, the mean value of **Infosys** stock returns appeared to be .0467 which exhibit that on the investment of shareholder, company is yielding market returns of around .04% which is quite less in comparison of TCS stock returns. The variation in stock returns of Infosys lies with minimum of -23.9013 and maximum of 15.5149. The minimum and maximum figures of Infosys stock returns indicate the shareholder's lost and gain percentage on each share. The standard deviation of 1.9107 reflects very less variation in returns.

Trend Analysis of TCS Stock Prices



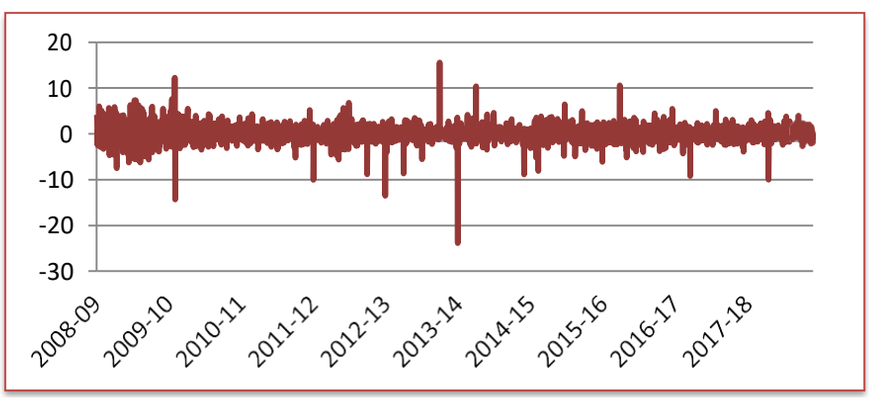
Graph-1: Trend Analysis of TCS Stock Prices

Above figure depicts the time series graph for the stock prices of Infosys. As per the chart, the stock price started at Rs 185 level, and then it reaches to a high of Rs 236.64 at 15th May'08. In the next 6 months, there was a decline in stock prices to Rs 174 level. After that in the next 2 years, stock rises to Rs 418 in December'10 and from here it went into consolidation till August'13. After this, bulls started running again and mark an all-time high of 623.75 on 23 rd May'16. There is an increasing trend in stock with some amount of consolidation.

UNIT ROOT TEST

So, it can be observed that time series of Infosys is non stationary. If the series is not stationary then it becomes essential to conduct some transformation steps. For this purpose, initially the stock prices of Infosys were converted into log normal stock returns and after that the test of Augmented- Dickey Fuller was applied on the series.

The results of ADF test which has been employed on the log normal stock return series of Infosys depicts the ADF value as -38.65942 and probability value as 0.00 which is less than .05. This ensures the stationarity of series at first differencing. Along with this, the value of Durbin-Watson stat i.e. 2 ensures that there exist no autocorrelation in the series.



Graph-3: Infosys Stationarity Series

Above figure depicts that the daily returns of Infosys between 1 April 2008 till 31 March 2018 except one major downfall, do not reveal any particular pattern means they tend to fluctuate randomly around zero. This leads to the conclusion that returns are largely independent of each other.

MODEL IDENTIFICATION

After achieving the results of stationarity in the series through log normal stock returns values, researcher stepped further and introduced Box-Jenkins methodology. The initial step in the process is to identify an appropriate model. In order to locate the best fitting ARIMA model for the stock returns of Infosys, a function "auto.arima" has been applied in R Studio. After applying the function best model has been estimated which identified the number of AR and MA terms on which returns of Infosys depends. Finally, the ARIMA model (3,0,1) came out to be as the best fit model for prediction of Infosys stock returns. In the estimated model, the value of AR comes out to be 3 which explains that stock returns of Infosys can be forecasted by considering stock returns of previous three days whereas I stands as 0 which shows stationarity of log normal returns series at differencing. Moreover, the value of MA comes out to be 1 which represents that stock returns of Infosys are affected by error term of previous one day.

MODEL ESTIMATION

Using R studio, following are the estimated parameters on the basis of model identified:

The coefficients depict the AR and MR terms of the ARIMA model whereas S.E shows the standard error. The mathematical equation of ARIMA model is

$$Y_t = 0.0464 - 0.7486\, Y_{t-1} - 0.0921\, Y_{t-2} - 0.1100\, Y_{t-3} + 0.0343 + 0.7536\, \varepsilon_{t-1}$$

Here the p-value of each coefficient has been examined in order to determine whether or not the parameters are significantly significant. As a result, it has been discovered that p-values of coefficients ar1, ar2, ar3, ma1 are found to be significant as its value comes out to be less than 5% significance level.

RESIDUALS DIAGNOSTIC

For a best forecasting model it is mandatory to diagnose the leftover residuals generated from the model. If these are left unchecked then, it may lead to the problem of autocorrelation. Therefore, the assumption of presence of autocorrelation among the residuals has been diagnosed by preparing correlograms of both autocorrelation as well as partial autocorrelation.

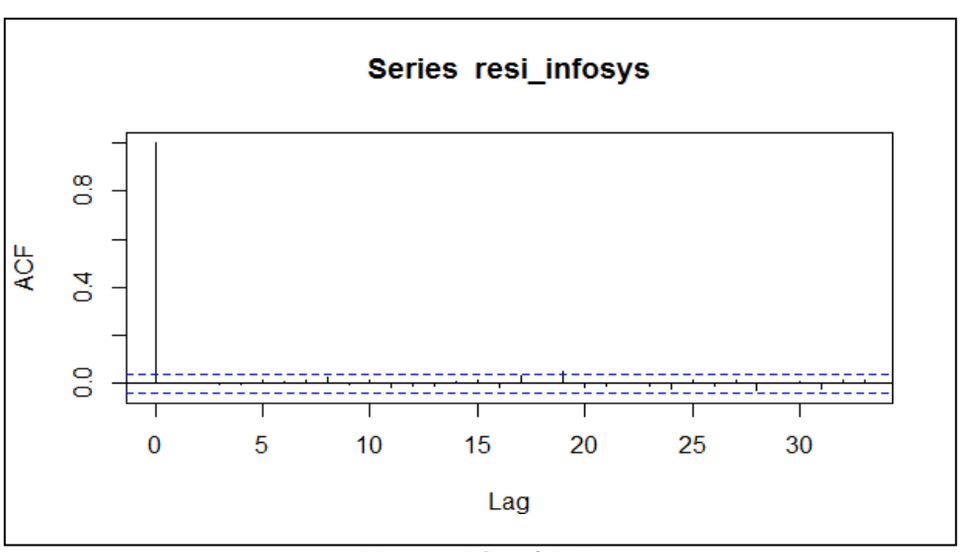


Figure-1: Plot of ACF

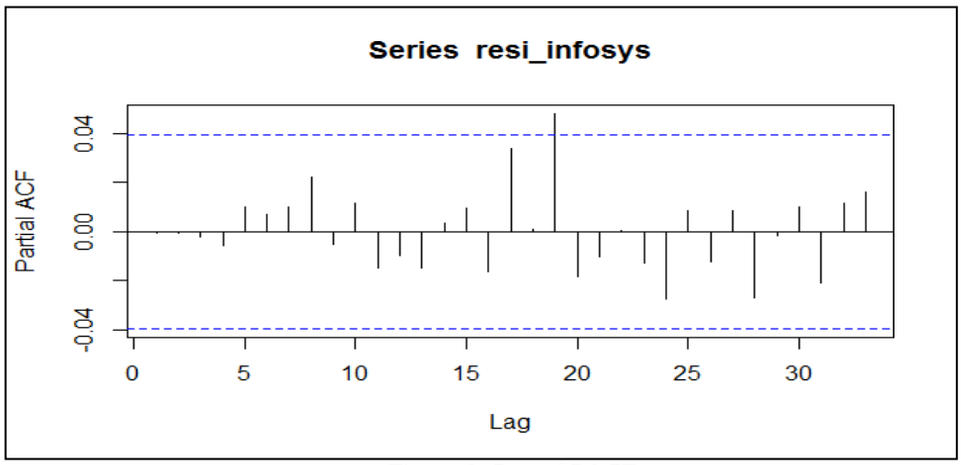


Figure-2: Plot of PACF

An analysis of above figure of ACF turned out that the error terms were not serially correlated as most of the spikes are not statistically significant. Moreover, from the plot of PACF , it can be determined that at lag 19 the spikes are statistically significant. This suggests that Infosys returns are correlated at lag 19 whereas remaining all lags are between the lines which suggests that there exists no autocorrelation in the error terms.

NORMALITY TEST

The second assumption of normality of residuals was satisfied by using Jarque-Bera test. This test is named after Carlos Jarque and Anil K. Bera. If it is far from zero, than it signals that the data do not have a normal distribution. The following is the assumption to check the normality of residuals.

Ho : Series is normally distributed
H1 : Series is not normally distributed

X- Squared	Df	p-value
29773	2	<2.2e-16

Table-2: Results of Jarque-Bera Test

Source: Output from R Studio

The p-value of Jarque-Bera test came out to be <2.2e-16 which is less than .05 which leads to not acceptance of null hypothesis and thus concludes that data is not normally distributed.

FORECASTING

Finally, when researcher have identified a successful ARIMA model (3, 0, 1), then forecasting of next 7 days log normal returns of Infosys has been done. The daily stock returns are compared with the returns generated by the best fit ARIMA (3,0,1) model which are shown in table below along with the calculated value of S.E.

Date	Actual Returns %	Forecast Returns %	SE%
2 April 2018	0.272899505	1.584918273	-1.31202
3 April 2018	0.224802076	0.953939177	-0.72914
4 April 2018	-1.368904456	0.725906964	-2.09481
5 April 2018	1.801278896	0.607885576	1.193393
6 April 2018	-1.380666526	0.532529249	-1.9132
9 April 2018	-1.637346534	0.479014842	-2.11636
10 April 2018	0.287625135	0.438492672	-0.15087

Source: Output from R Studio

Table-3: Forecasted Returns of Infosys

Above table represents actual stock returns and forecasted stock returns of Infosys along with its standard error. The results depicts that only on 5 April 2018, the value was under forecasted by 1.19%,

except this day remaining all the returns from 2nd April 2018 till 10th April 2018 have been over forecasted . The range of over forecasted values lies 0.15% to 2.11% which confirms that model is precise or best fit for prediction.

FINDINGS & CONCLUSION

The findings of the study from the above analysis states that ARIMA model (3,0,1) is the best fitted model to forecast the log normal stock returns of Infosys. The results represented that only on 5 April 2018, the value was under forecasted by 1.19%, except this day remaining all the returns from 2nd April 2018 till 10th April 2018 have been over forecasted. The range of over forecasted values lies between 0.15% to 2.11% which confirms that model is precise or best fit for prediction.

From the above findings it can be concluded that ARIMA model has sufficient potential to predict future values in short run. The implication of the study is that it is expected to be worthy for prospective investors by guiding them to invest or disinvest in a particular stock at correct time.

BIBLIOGRAPHY

Afeef, M; Ihsan, A; & Zada, H. (2018). Forecasting Stock Prices through Univariate ARIMA Modeling. NUML International Journal of Business & Management, 13(2), 130-143.

Asteriou, D., & Hall, S. G. (2007). Applied Econometrics: a modern approach, revised edition. Hampshire: Palgrave Macmillan.

Devi, B. U; Sundar, D; & Alli, P. (2013). An effective time series analysis for stock trend prediction using ARIMA model for nifty midcap-50. International Journal of Data Mining & Knowledge Management Process, 3(1), 65.

Gay, R. D. (2016). Effect of macroeconomic variables on stock market returns for four emerging economies: Brazil, Russia, India, and China. International Business & Economics Research Journal (IBER), 15(3), 119-126.

Gujarati, D., & Porter, D. C. (2004). Basic Econometrics, 2004. Editura McGraw-Hill, 858.

Gupta, S., & Kashyap, S. (2015). Forecasting inflation in G-7 countries: an application of artificial neural network. Foresight, 17(1), 63-73.

Hamjah, M. A. (2014). Forecasting major fruit crops productions in Bangladesh using Box-Jenkins ARIMA model. prospects, 5(7).

Ivanovic, Z; Bogdan, S; & Baresa, S. (2013). Forecasting Croatian stock market index: Crobex. UTMS Journal of Economics, 4(2), 79-91.

Malik, F; Wang, F; & Naseem, M. A. (2017). Econometric estimation of banking stocks. The Journal of Developing Areas, 51(4), 207-237.

Mondal, P; Shit, L; & Goswami, S. (2014). Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices. International Journal of Computer Science, Engineering and Applications, 4(2), 13.

***