International Journal of Economic Perspectives, 2010, Volume 4, Issue 4, 617-622.

Forecasting Stock Market Volatility in India – Using Linear and Non - Linear Models

Kaliyaperumal SRINIVASAN, Ph.D
Department of Management Studies, Christ University, Bangalore, Karnataka - 560 029, INDIA. Tel: +91-96860-59109; +91-9488818116, E-mail: ksrinivasan1979@gmail.com, Alternative Email: srinivasan@christuniveristy.in.

Malabika DEO, Ph.D
Department of Commerce (SOM), Pondicherry Central University, Kalapet, Puducherry, 605 014, INDIA. Tel: 0413-2340755; +91-9442140745, E-mail: malavika_deo@yahoo.co.in.

ABSTRACT

Volatility models and their forecasting performance attracted the interest of many economic agents, especially for financial risk management. The role of economic agents is to decide which one will be best model for forecasting volatility. This paper examines the modeling and forecasting performance of BSE Sensex daily stock market returns over the period from 1 July 1997 to 31 October 2008, by using simple Random Walk, GARCH, EGARCH and TGARCH models. The out-of-sample forecasts are evaluated by using MAE, RMSE, MAPE and Theil – U Statistics. The result suggests the standardized residual of white noise series strongly rejects the null hypothesis for GARCH model and capture the serial dependence and inherent nonlinearity series. Moreover, Random walk model dominates the forecasting performance and it is considered as the best model followed by the TGARCH model.

JEL Classification: C32; C53; G15.

Key Words: BSE Sensex; Emerging Markets; GARCH Model; Volatility; Forecasting.

1. INTRODUCTION

Extensive work has been done on traditional econometric time series for modelling the conditional mean of random variables. However, many theoretical and empirical studies are designed to work with the conditional variance in developed markets Dimson and Marsh (1990), McMillan, Speight and Gwilym (2000). Stock prices volatility in financial markets has received a great attention from academics, policy makers and practitioners over the past decades because it can be used as a measure of risk. Modelling and forecasting the conditional volatility of return was highlighted on asset pricing, option pricing, and value at risk and hedging strategies by Jondeau and Rockinger (2003).

Financial time series returns often exhibit some well-known characteristics. First, large changes tend to be followed by large changes and small changes tend to be followed by small changes, which mean that volatility clustering is observed in financial returns data. Secondly, financial time series data often exhibit leptokurtosis, which indicate that the return distribution is fat-tailed as observed by Mandelbrot (1963), Fama (1965), Laurent and Peters (2002). Moreover, changes in stock prices tend to be negatively related to changes in stock volatility which is identified to be "leverage effect" Black (1976), Christie (1982), Nelson (1991), Koutmas and Saidi (1995). Recently, several authors have investigated the volatility of stock market by applying econometric models and suggested that, no single model is superior Akgiray (1989), Pagan and Schwert (1989). Brailsford and Faff (1996) and Koutmos (1998) examine the predictive performance of several statistical methods with GARCH and TGARCH models for Australian stock exchange. Dimson and Marsh (1990) examine various technical methods of predicting the volatility of UK stock market returns and find that exponential smoothening and regression model performed best according to their criteria. There is a large literature on modelling and forecasting volatility, however only a limited study has appeared in the literature focusing on the Indian stock market. Varma (1999) examined the volatility estimation models comparing GARCH and EWMA models in the risk management setting. Pandey (2002) analyzed the extreme value estimators and found the performance with Parkinson estimator for forecasting volatility over these horizons. Kumar (2006) examined the comparative performance of volatility forecasting models in Indian markets and the results were found contrary to Brailsford and Faff (1996). However, further research is needed to forecast the volatility of Indian stock market for an in-depth understanding the characteristics of Indian markets.

This study can be considered as one of the few attempts made to model the most prominent features of the time series of BSE sensex such as volatility clustering, excess kurtosis, and fat tailed by applying the most popular techniques proposed by Engle (1982). To capture the above characteristics, ARCH class of models were introduced by Engle (1982) and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) by Bollerslev (1986) and Taylor (1986). Since the intrinsically symmetric GARCH model does not cope with the asymmetry issues or so called leverage effect, the Exponential Generalized Autoregressive Conditional Heteroskedasticity process (EGARCH) by Nelson (1990) is suggested. Finally, to capture asymmetries in terms of negative and positive shocks TGARCH (Threshold Generalized Autoregressive Conditional Heteroskedasticity) model was introduced by Zakoian (1990) and Glosten, Jaganathan and Runkle (1993).

The structure of this paper is organized as follows: Section 2 describes brief discussion about the methodology, while Section 3 incorporates the data used and provides empirical results of the models. Section 4 details the forecasting performance of the estimated models. Section 5 ends with some concluding remarks.

2. BRIEF DISCUSSIO O N THE METHODOLOGY

Random Walk Model

The random walk model is the simplest possible models, where the Ordinary Least Square (OLS) method are constructed on the assumption of constant variance. As per, efficient market hypothesis the competing market participants reflect information instantly hence are useless in predicting future prices. The basic model for estimating stock returns fluctuation by using OLS in the naïve random walk model is given below:

$$R_t = \mu + \varepsilon_t$$

where, μ is the mean value of the returns, it is expected to be insignificantly differ from zero; and ε_t, the error term should not be serially correlated over time.

GARCH

Bollerslev (1986) extended Engle's ARCH model to the GARCH model and it is based on the assumption that forecasts of time varying variance depend on the lagged variance of the asset. An unexpected increase or decrease in returns at time t will generate an increase in the expected variability in the next period. The basic and most widespread model GARCH can be expressed as;

$$R_t = a + bR_{t-1} + \varepsilon_t$$

$$\varepsilon_t \mid I_{t-1} \; N(0, h_t),$$

$$h_t = \alpha_0 + \sum_{i=1}^{p} \beta_i h_{t-1} + \sum_{i=1}^{p} \lambda_j u_{t-j}^2$$

where, $\lambda_j > 0$, $\beta_i \geq 0$. The GARCH is weekly stationery $\Sigma\beta_i + \Sigma\lambda_j < 1$, the latter two quantifying the persistence of shocks to volatility Nelson (1992).

In particular, volatility forecast are increased following a large positive and negative index return, the GARCH specification that capturing the well-documented volatility clustering evident in financial returns date Engle (1982).

TGARCH

In TGARCH model, it has been observed that positive and negative shocks of equal magnitude have a different impact on stock market volatility, which may be attributed to a " leverage effect" (Black, 1976). In the same sense, negative shocks are followed by higher volatility than positive shocks of the same magnitude (Engle and Ng, 1993). The threshold GARCH model was introduced by the works of Zakoian (1990) and Glosten, Jaganathan and Runkle (1993). The main target of this model is to capture asymmetric in terms of negative and positive shocks and adds multiplicative dummy variable to check whether there is statistically significant different when shocks are negative. The conditional variance for the simple TGARCH model is defined by;

$$R_t = a + bR_{t-1} + \varepsilon_t$$

$$\varepsilon_t \mid I_{t-1} \ N(0, h_t),$$

$$h_t = \alpha_0 + \sum_{i=1}^{p} \beta_i u_{t-1}^2 + \sum_{j=1}^{q} \lambda_j h_{t-j} + \delta u_{t-1}^2 d_{t-1}$$

where, d_t takes the value of 1 if ε_t is negative, and 0 otherwise. So "good news" and "bad news" have a different impact.

EGARCH

The Exponential GARCH model specifies conditional variance in logarithmic form, which means that there is no need to impose estimation constraints in order to avoid negative variance Nelson (1991). The mean and variance equation for this model is given by;

$$R_t = a + bR_{t-1} + \varepsilon_t$$

$$\varepsilon_t \mid I_{t-1} \ N(0, h_t),$$

$$\log(h_t) = \alpha + \sum_{j=1}^{q} \beta_j \left| \frac{u_{t-j}}{\sqrt{h_{t-j}}} \right| + \sum_{j=1}^{q} \lambda_j \frac{u_{t-j}}{\sqrt{h_{t-j}}} + \sum_{i=1}^{p} \delta_i h_{t-1}$$

where, δ captures the asymmetric effect. The exponential nature of EGARCH ensures that the conditional variance is always positive even if the parameter values are negative; thus there is no need for parameter restrictions to impose non-negativity.

3. DATA AND EMPIRICAL RESULTS

Data and Descriptive Statistics

The BSE Sensex index was first compiled in 1986 and represents stock prices of 30 largest stocks representing large, well-established and financially sound companies across key sectors. The total sample spans the period from July 1, 1997 to October 31, 2008 for a total of 2805 daily stock return observations of BSE Sensex. The full sample is categorized into two parts; an in-sample and out-sample. In- sample consists of 2560 observation from 1 July 1997 to 31 October 2007 in order to estimate the parameters of each model. An out of sample composed of 246 observations from 1 November 2007 until 31 October 2008, in order to make forecast. Returns are defined as the natural logarithm of price relatives; that is, $r_t = \log X_t / X_{t-1}$ is the daily capital index. The below table displays some descriptive statistics about the sensex indices;

A Summary of some characteristics of the r_t series is given in Table: 1. The mean and standard deviation of the daily return are quite small and close to zero, this indicate that the fluctuations of the BSE sensex is quite normal. There is also evidence of positive skewness and kurtosis in sensex indices, which is a confirmation of the stylized fact related to fat tails and extreme values with high frequencies data. Moreover, the kurtosis exceeds 3, which is the normal value of the positive skewness mean that the right tails is particularly extreme. As a result, the Jarque-Bera test rejects the null of the normality. The Ljung-Box Q statistics for the return series observes that the index shows evidence of ARCH effects judging from the significant autocorrelation coefficients. The conventional ADF (Augmented Dickey Fuller) and Phillips Perron (PP) statistics are -7.48808 and -57.0139 respectively. The results allow us to reject the null hypothesis that returns have unit root in favour of alternative hypothesis of stationery (even at 1 per cent Mac Kinnon critical value).

Estimated Results

The parameter estimates for typical and parsimonious Random Walk (RW), GARCH, TGARCH and EGARCH models for daily BSE Sensex stock return series used the robust method of Bollerslev – Wooldridge's quasi maximum likelihood estimator assuming the Gaussian standard normal distribution. Next we use a combination of information log-likelihood (LL) values and a set of model diagnostic tests (ARCH - LM test, Q – statistics and BDS test) to choose the volatility models that best models conditional variance of the BSE Sensex. For this exercise, the nonparametric BDS test for serial independence is applied Chris Brooks (2002).

The main results of Random Walk, GARCH, TGARCH and EGARCH models are presented in Table: 2 during the period from 1st July 1997 to October 31, 2008. In the mean equation the coefficients are statistically significant at 1 per cent level, but in EGARCH models the estimated parameter indicated insignificant effect. The result suggests the mean (μ) return series is significantly different from zero, which is inconsistent with the random walk hypothesis. In variance equation the ARCH and GARCH effect are positive for GARCH and EGARCH model and very close to one, indicating that volatility shocks are quite persistent, but for TGARCH model envisage negative trend. Moreover, the battery of the diagnostic tests applied to the residuals of RW model is statistically significant at 1 per cent level. ARCH – LM test and Q – Statistics of the standardized residuals shows the presence of significant ARCH effects and autocorrelation function for all the four models. Overall, using the maximum AIC and maximum LL models, the preferred model is the TGARCH model.

BDS Diagnostic Tests

This test can be applied to a series of estimated residuals to check whether the residuals are independent and identically distributed (iid). The BDS test is a portmanteau test for time based dependence in a series. It can be used for testing against a variety of possible deviations from independence including linear dependence, non-linear dependence, or chaos. Brock, Dechert, Scheinkman and LeBaron (1996). According to the principle, once any linear or non-linear structure is removed from the data the remaining structure should be due to an unknown non-linear data generating process. In this case, if a model is the true data generating process then we expect its residuals to be white noise otherwise we reject the null hypothesis and consider the model as inadequate to capture all of the relevant features of the data. Hence the BDS test statistic for the standardized residuals will be statistically significant.

In Table 3 the ordinary residuals of Random Walk model and standardized residuals of GARCH, TGARCH and EGARCH models were used for the BDS test. Bootstraps with 1500 new samples were applied for all the four models. The results of the hypothesis indicate both the bootstraps and asymptotic p-values of BDS test suggest absence of nonlinear dependence in BSE Sensex. Furthermore, the standardized residual of the white noise series strongly rejects the null hypothesis at 1 per cent level of significance for GARCH model and capture the serial dependence and inherent nonlinearity series.

4. FORECAST EVALUATION

In order to evaluate the forecasting performance of different models we use four forecast error statistics, namely, the mean absolute error (MAE), the root mean square error (RMSE), the mean absolute percentage error (MAPE) and the Theil- U statistics, defined as follows :

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | \hat{\sigma}_i - \sigma_i |$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\hat{\sigma}_i - \sigma_i)^2}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} | (\hat{\sigma}_i - \sigma_i) / \sigma_i |$$

$$TU = \frac{\sum_{i=1}^{n} (\sigma_{i-1} - \sigma_i)^2}{\sum_{i=1}^{n} (\sigma_{i-1} - \sigma_i)^2}$$

Where in all the above statistics *'n'* stand for the number of out of sample forecasts. MAE and RMSE are two of the most popular measures to evaluate the forecasting capability of a model. MAE measures the average absolute forecast error. It is a conventional forecast accuracy evaluation criterion which does not permit the offsetting effects of over prediction and under prediction. Whereas RMSE weights greater forecast errors more heavily in the average forecast error penalty. Another criterion which is popular among financial researchers is 'Theil – U' statistics (Theil, 1996). Theil- U statistics uses the random walk model as a benchmark model for comparing the level of accuracy of forecast model. A 'U' statistic of one indicates that the model under consideration and the benchmark model have equal forecasting capability. While if 'U < 1 the forecasting model under consideration is better than the benchmark model and vice versa for 'U' > 1. In short, the model that exhibits the lowest values of the error measurement technique is considered to be the best model.

The results reported in Table 4 shows that the random walk model has outperformed all the other models and provides the most accurate forecast. On the basis of MAE, RMSE forecast error statistics and Theil's – U statistics. Random walk model dominates the forecasting performance and it is considered as the best model followed by the TGARCH model. On the other hand the GARCH model is the worst performing model under the same criterion. However, under the MAPE forecast error statistics the GARCH model is considered as the best model and the EGARCH is ranked second best. Despite its mathematical and statistical simplicity, the random walk model provides the most accurate forecast compared to other competing models in the study. Our finding that random walk model is the best model and offers comparatively high forecast accuracy is consistent with the study of Stock and Waston (1998) which argues in favour of random walk model among other rival models for the US macroeconomics series.

5. CONCLUSION

Volatility forecasting is a widely researched area in the finance literature. The performance of forecasting models of varying complexity has been investigated according to a range of measures and generally mixed results have been recorded. On the one hand, some argue that relatively simple forecasting techniques are superior, while other suggests complex ARCH - type models worthwhile. In our analysis, a range of forecasting models like Random Walk model, GARCH, TGARCH and EGARCH model. Furthermore, we compare the forecasting techniques based on both symmetric and asymmetric error statistics are compared.

This paper empirically identifies some of the key volatility characteristics of BSE Sensex daily stock market returns over the period from 1 July 1997 to 31 October 2008. The results suggest that volatility can be best specified as a process of conditional heteroskedasticity in Indian stock market. First, the results of ARCH and GARCH effect are positive for GARCH and EGARCH model and close to one, indicate that volatility shocks are quite persistent, but for TGARCH model the observed result are negative trend and performed the best. Secondly, the battery of diagnostic tests applied to the standardized residual of white noise series strongly rejects the null hypothesis for GARCH model and capture the serial dependence and inherent nonlinearity series. Finally, Random walk model dominates the forecasting performance and it is considered to be the best model followed by TGARCH model. These findings are contrary to the findings of Brailsord and Faff (1996) who found no single method as superior. By and large, we came to the important conclusion that reducing the sample period for estimation improves the accuracy of predicting future observations of the BSE Sensex stock returns.

REFERENCES

Akgiray. V. (1989), Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts, *Journal of Business*, 62: 55-80.

Black, F. (1976), Studies of Stock Market Volatility Changes, *Proceedings of the American Statistical Association*, Business and Economic Statistics Section, 177-181.

Bollerslev, T. (1986), Generalised Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31: 307-27.

Brailsford T.J. and R.W. Faff, (1996), an Evaluation of Volatility Forecasting Techniques, *Journal of Banking and Finance*, 20: 419-38.

Brock, W.A., Dechert, W., and Scheinkman, H and LeBaron, B. (1996), a Test for Independence Based on the Correlation Dimension, *Econometric Reviews*, 15: 197-235.

Brooks, C. (2002), *Introductory Econometrics for Finance*, 2nd Edition, Cambridge University Press, Cambridge.

Christie, A. A. (1982), the Stochastic Behaviour of Common Stock Variance Value, Leverage and Interest Rate Effects, *Journal of Financial Economics*, 10: 407-32.

Dimson, E. and Marsh, P. (1990), Volatility Forecasting without Data-Snooping, *Journal of Banking and Finance*, 14: 399 – 421.

Engle, R. F. (1982), Autoregressive Conditional Heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50: 987-1007.

Engle, R. F. and Ng, V.K. (1993), Measuring and Testing the Impact Of News On Volatility, *Journal of Finance*, 48: 1749–78.

Fama, E., (1965), the Behaviour of Stock Market Prices, *Journal of Business*, 38 (1): 34-105.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993), On the Relation between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks, *Journal of Finance*, 48 (5): 1779-91.

Jondeau, E. and Rockinger, M. (2003), Conditional Volatility, Skewness and Kurtosis: Existence, Persistence, and Co-movements, *Journal of Economic Dynamic and Control*, 27: 1699 – 737.

Koutmos, G. and Saidi, R. (1995), Asymmetric Volatility and Risk Return Tradeoff in Foreign Stock Markets, *Journal of Multinational Financial Management*, 2: 27-43.

Koutmos, G. (1998), Asymmetries in the Conditional Mean and the Conditional Variance: Evidence from Nine Stock Markets, *Journal of Economics and Business*, 50: 277-90.

Kumar, S. S. (2006), Comparative Performance of Volatility Forecasting Models in Indian Markets, *Decision*, 33 (2): 26 – 40.

Laurent, S. and Peters, J-P, (2002), GARCH 2.2: An Ox Package for Estimating and Forecasting Various ARCH models, *Journal of Economic Surveys*, 16: 447–85.

Mandelbrot, B. (1963), the Variation of Certain Speculative Prices, *Journal of Business*, 36: 394-419.

McMillan, D., Speight, A. and Gwilym, O. (2000), Forecasting UK Stock Market Volatility, *Applied Financial Economics*, 10: 435-48.

Najand, M. (2002), Forecasting Stock Index Futures price Volatility: Linear Versus Nonlinear Models, *the Financial Review*, 37: 93 – 104.

Nelosn, D. B. (1991), Conditional Heteroskedasticity in Asset Returns: a New Approach, *Econometrica*, 55: 703 –8.

Pagan, A. R. and Schwert, G. W. (1989), Alternative Models for Conditional Stock Volatilities, *Journal of Econometrics*, 45: 267-90.

Pandey, A. (2002), the Extreme Value Volatility Estimators and Their Empirical Performance in Indian Capital Markets, NSE Working Paper No. 52 (Downloadable from www.nseindia.com).

Stock, J. and Watson, M. (1998), a Comparison of Linear and Non-Linear Univariate Models for Forecasting Macroeconomic Time Series, NBER Working Paper, No. 6607.

Taylor, S. J (1986), *Modelling Financial Time Series*, Chichester, John Wiley Publications.

Varma, J. R (1999), Value at Risk Models in the Indian Stock Market, IIMA Working Paper No: 99 – 07 – 05, India.

Theil, H. (1996), *Applied Economic Forecasting*, North – Holland, Amsterdam.

Zakoian, J. (1994), Threshold Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Economic Dynamics and Control*, 18: 931 – 55.