Volatility Estimation In The Indian Stock Market Using Heteroscedastic Models

** M.P. Rajan*

INTRODUCTION

The investment is an art of science. If the resources and benefits taken in the form of money, investment is the present commitment of money for the purpose of making money in the future. Investment science is the application of scientific tools to investments. The scientific tools primarily used vary from simple to a higher level of mathematics. The art to investment is knowing what to analyze and how to go about it. The analysis involves examining alternatives and deciding which alternative is the most preferable. The analysis looks similar to other decision problems in practice, but it differs to it with respect to the fact that most investments are carried out within the framework of the financial market and these markets provide alternatives not found in other decision situations and hence, investment analysis is unique and unusually a powerful methodology. The investment in stock market involves risk as the market return is very uncertain. In order to capture this uncertainty, one needs to employ sophisticated mathematical models.

India is considered to be an emerging economic powerhouse in the global market. The Indian market also exhibits ups and downs, which is unprecedented due to global recession. The rapid pace change and volatility that have buffeted the global economy over the past few months had an impact on the Indian economy too. Even then, the investors looked at the Indian Stock market with great hope and optimism. An investor is very much concerned about the return he makes from the purchase of an asset. The performance of the stock market and stock price changes over time is not only a common concern of an institutional investor, but for each individual. In India, the price indices such as SENSEX and NIFTY give good measure of price movements over time. But in any financial market, whether it is a stock price or commodity price or exchange rate, it behaves in a particular fashion. Many empirical studies revealed that the financial market exhibits common features such as Fat tails; Asymmetry; Aggregate normality; Absence of serial correlation; Volatility clustering; Time varying cross-correlation and these features commonly known as stylized facts. One of the fundamental concerns about the stock market is the market volatility. Volatility of an asset in simple terms can be defined as the standard deviation of the asset return. Economists believe that volatility can be explained by efficient market hypotheses. In this paper, we are concerned with studying efficient market model for volatility that captures the Indian stock market behaviour.

Our idea is to model the statistical volatility rather than implied volatility. The statistical volatility depends on the choice of the statistical model that is applied to historical asset return data. The statistical model is usually a time series model. Applying the model to the historical data will generate the statistical estimates of volatility for the past where the historical data are available. It will also generate the forecasts of the volatility from now until some future point in time called the risk horizon. Unlike prices, the volatility is not directly observable in the market. They can only be estimated. Therefore, it is very important that a statistical volatility model provides more accurate estimate or forecast. Moreover, volatility forecasting is crucial for option pricing, risk management and portfolio management. The researcher has organized the paper in the following manner. In section 2, the researcher discusses various types of models to estimate the volatility in the financial market. In section 3, the researcher focuses on the applicability of such models in Indian context and section 4 covers the conclusion.

MODELING VOLATILITY

The volatility model can be of either a constant or time varying one. Constant volatility model usually refers to the unconditional volatility of the return process. It is a finite constant which is the same throughout the data generating process and can be calculated as the standard deviation of the unconditional distribution of the return process. An unconditional volatility is only defined if we assume that the return series is generated from a stationary stochastic

**Associate Professor In Mathematics and Finance*, Indian Institute of Science Education and Research Thiruvananthapuram, Thiruvananthapuram- 695 016, Kerala. E-mail: rajanmp@iisertvm.ac.in

process. But this is is not a reasonable assumption that is commonly encountered in the financial market. On the other hand, the time varying model describes a process for the conditional volatility. A conditional distribution, in this context, is a distribution that governs a relation at a particular instant in time and the conditional volatility at the time is the square root of the variance of the conditional distribution at time t. Moreover, a conditional forecast is considered to be superior to an unconditional forecast. Both conditional mean and variance could change at every time period throughout the process, but for the purpose of estimating and forecasting the conditional volatility, the conditional mean is often assumed to be constant. In the following sections, the researcher addresses the question of how to model the conditional volatility.

ARCH AND GARCH TYPE MODELS

In conventional econometric models, the variance of the disturbance term is assumed to be constant. However, many econometric time series do not have a constant mean and most exhibit the phases of relative tranquility followed by periods of high volatility. In such circumstances, the assumption of homoscedasticity (constant variance) is inappropriate. Moreover, the returns in many financial markets are not well modeled by an independent and identically distributed process. Except at high frequency, the return might not show any sign of autocorrelation, but quite often, a strong autocorrelation can be seen in squared returns. A positive autocorrelation in squared return indicates that financial market volatility comes in clusters, where tranquility periods of small returns are interspersed with volatile periods of large returns. ARCH(Auto Regressive Conditional Heteroscedastic) model was first introduced in literature by **Engle**[3] and later extensively studied by many researchers **(cf. Bollerslev**[1], **Bollerslev, Engle and Nelson**[2], **Higgins and Bera**[4]. In ARCH model, one would model the conditional variance of the asset return through maximum likelihood procedure. A general form of an ARCH(m) model is as follows:

$$r_t = \mu + \varepsilon_t \quad (1)$$

$$\varepsilon_t = \sigma_t w_t \quad (2)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i \varepsilon_{t-i}^2 \ , \quad \alpha_0 > 0; \ \alpha_i \geq 0, i = 1,2......m \quad (3)$$

where r_t is the return of the series; ε_t is the disturbance term; σ^2_t is the conditional volatility; w_t is the i.i.d white noise with mean 0 and variance 1. The conditions imposed on the parameters α_t ensure that the volatility is non-negative. Because of the volatility persistence in the financial market, it is found in practice that typically the order *m* is very large. The impact of large order is essentially responsible for high computational effort. To over come this difficulty, later, more advanced models such as GARCH (Generalized Auto Regressive Conditional Hetroscedastic) have been studied by Bollerslev[1]. This model incorporates additional dependencies of past values of σ^2_t in the equation (3). A general form of GRACH (p,q) model for the conditional volatility is:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \sigma_{t-j}^2 \ , \quad \alpha_0 > 0; \alpha_i \geq 0, i = 1,2.....p \ ; \beta_j \geq 0, j = 1,2......q \quad (4)$$

The conditions imposed on the parameter ensure that volatility is strictly positive. In practice, a GARCH(1,1) model performs better than a larger order ARCH model and many practitioners still favor a GARCH modeling for the volatility estimation. Because of this popularity, many variants of GARCH models also have been studied in literature **(cf. Nelson** [7]**)**. The predictive power of the model depends on how accurate one could estimate σ_t. The practice of choosing daily squared return as a proxy for conditional variance gives a wrong signal for model selection. In time series volatility modeling, (as is described above) is capable of forecasting the volatility based on the past observations, but they are not able to predict the shocks that are new to the system. Once the shocks enter the system, the merits of a model depend on how well it can capture the so-called stylized facts of the financial market. GARCH models could capture many characteristics of the stylized facts, but not completely. However, GARCH models cannot capture the leverage effect that is asymmetric in nature as they are basically symmetrical in nature. Another drawback of the GARCH modeling is that the conditional variance formulation of σ^2_t gives too much weight to the error caused by the shocks. To overcome this problem, a rescaling of the volatility can be employed through a logarithmic transformation. These thoughts lead to the development of more advanced GARCH models such as Exponential GARCH (EGARCH) **(cf. Nelson** [7]**)**, Threshold Garch (TGARCH) **(cf. Glosten, Jaganathan and Rukle [5]**, **Zakonian[8])** in literature. The EGARCH (p,q) model has the representation:

$$\log\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \frac{\|\varepsilon_{t-i} + \gamma_i \varepsilon_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{q} \beta_j \log\sigma_{t-j}^2 \quad \textbf{(5)}$$

and TGARCH(p,q) model has the form:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} (\alpha_i + \gamma_i S_{t-i})\varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \quad \textbf{(6)}$$

where,

$$S_{t-i} = \begin{cases} 1, if\ \varepsilon_{t-i} < 0 \\ 0, if\ \varepsilon_{t-i} \geq 0 \end{cases}$$

APPLICATIONS OF MODELS IN THE INDIAN MARKET

In this section, the researcher investigates the applicability of the models in the Indian market. Most of these models have been widely used in developed markets, only few works have attempted to study application of these models in the Indian context. One such study is the work of **Kumar** [6]. Kumar studied basic models and GARCH model and the study indicates that GARCH model performs better than the basic models. As discussed in the previous section, the simple variance of the past return gives a measure of the volatility, but it will not give the correct market volatility. There are other basic models such as Extreme-value estimators, Exponential Weighted Moving Average(EWMA) discussed in literature, but the researcher mainly concentrates on the GARCH-type models discussed here. Another reason to start with GARCH model for the discussion is that it is widely considered as practitioners model. The researchers took SENSEX as a proxy for the stock market and collected the daily data from BSE site for a period ranging from July 1,1996 to July 3,2009. It is plotted in the Figure 1 and the log return of the data is also plotted in the Figure 2. The researchers considered only equity market and the models can be applied in other markets as well.

Figure 1: Sensex Data

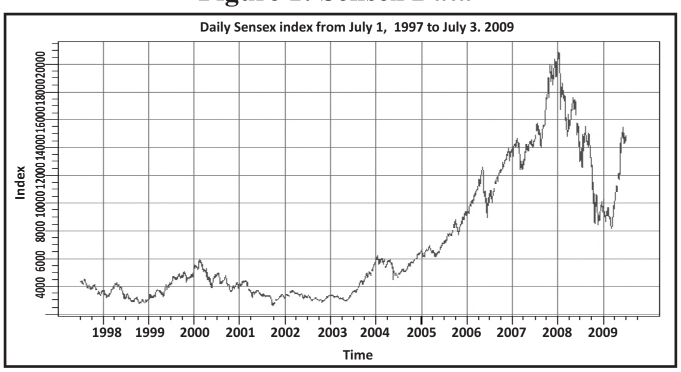
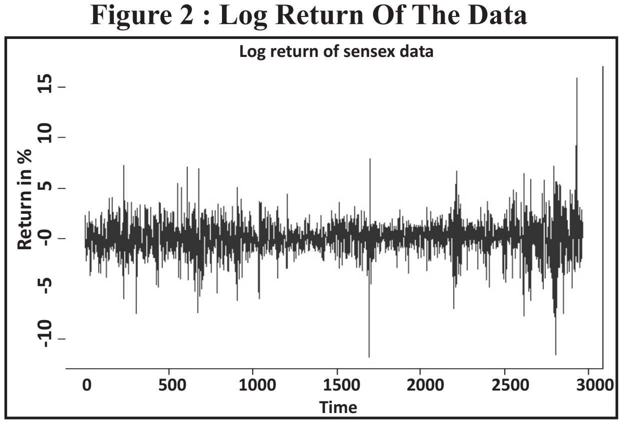


Figure 2 : Log Return Of The Data



Summary statistics of the data is presented in Table 1. It indicates that the average daily return is *0.04%* and the standard deviation of the return series is *1.8%*. The coefficient of skewness is negative and it indicates negative asymmetry. The coefficient of kurtosis is positive and is significantly higher and more than 3. It indicates that return is leptokurtic compared to normal distribution.

Table 1: Summary Statistics

Minimum	1Quartile	Median	3Quartile	Max	Mean	Std	Skewness	Kurtosis
-0.118	-0.0089	0.001164	0.01025	0.1599	0.00041	0.018	-0.09165	8.017

Test for normality and autocorrelation is presented in Table 2.

Table 2: Test for Normality and Autocorrelation

Test for Normality: Jarque-Bera				Test for Autocorrelation:-Ljung-Box			
Null Hypothesis: Data is normal				Null Hypothesis: No Autocorrelation			
Statistics	P-Value	Distribution	DF	Statistics	PValue	Distribution.	DF
3113.06	0.0000	Chi-Sq	2	51.7126	0.0000	Chi-Sq	10

Jarque-Bera test for normality comes out with P-value as 0 with the null hypothesis as *"return series is normal"*. Since the P-value is zero less than *5%* level, we simply reject the null hypothesis that return series is normal. This in fact justified by the Q-Q plot (see Figure 3), which shows that data deviates from the normal line. The Ljung-Box test for autocorrelation comes out with P-value as zero suggest that the null hypothesis that *"no auto-correlation"* can be rejected. This is further supported with the ACF analysis. Even though the ACF of the return series shows mild autocorrelation, the ACF of squared returns exhibit significant auto correlation. Since the squared return exhibits second order moment of the original return, the result shows that variance of the data conditional on its past history might change overtime. In other words, the return data signals the conditional heteroscedasticity feature. The augmented Dicky-fuller Test(ADF) with null hypothesis as *"there is a unit root"* came out with a p-value of *1* suggests that the we cannot reject the null hypothesis and presence of unit root is ruled out indicating that the log return series are stationary.

Figure 3: Q-Q Plot

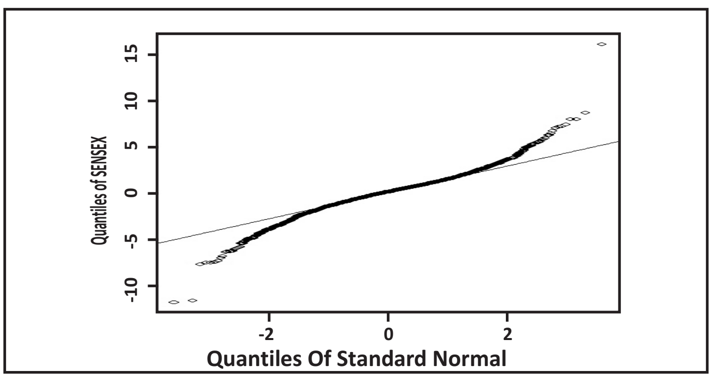


Figure 4: ACF of the return

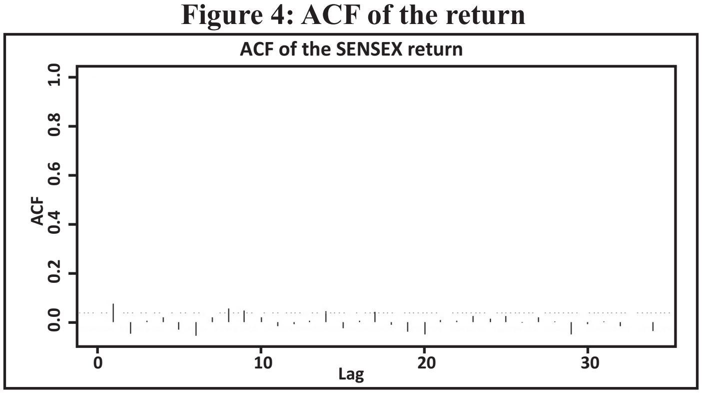
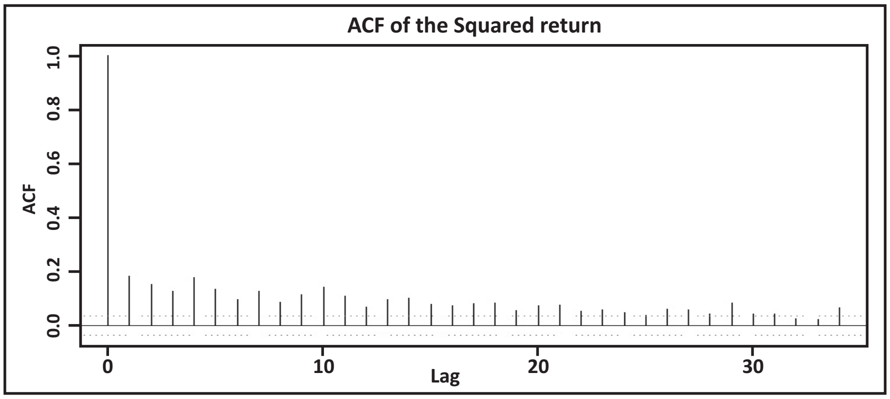


Figure 5: ACF of Square Return



ESTIMATION USING GARCH, EGARCH AND TGARCH MODELS

In order to estimate the volatility using the GARCH type models, we follow the following steps:

1. First we make sure that there is an ARCH effect in the return series.

2. If there is an arch effect, we would use the GARCH /EGARCH /TGARCH model for modeling the volatility.

3. We perform a residual analysis to check the model fitted the data well.

Step 1 is achieved by first applying an AR(p) model to the data set and then obtain the square of the fitted error, ε^2_t. Subsequently, we model the squared error in terms of q lagged variables and apply a regression. Then we employ a simple Lagrangian Test for the NH: the regression coefficients, $a_t = 0, i = 1, ...q$ with the LM statistics as, $LM = TR^2 \sim \chi_q^2$ where T is the sample size and R^2, the R-square of the regression. The test result is presented in Table 3. Since the P-value of the test is zero, we conclude that we can reject the hypothesis that there is no arch effect.

Table 3: Test for Arch Effects

Test for Arch Effects: LM Test			
Null Hypothesis: No Arch Effect			
Statistics	P-Value	Distribution	DF
277.555	0.0000	Ch-Square	20

Table 4: Estimated Results

Model	Coeffs.	Value	Stan.Error	t-value	Prob.	AIC	BIC
Garch(1,1)	μ	0.1442	0.023107	6.238	5.048e-10	11151.71	11175.69
	α_0	0.0807	0.010577	7.629	3.153e-14		
	α_1	0.1353	0.009420	14.363	0.000e+00		
	β_1	0.8453	0.009745	86.745	0.000e+00		
EGarch(1,1)	μ	0.1398	0.020462	6.832	1.015e-11	11163.3	11187.28
	α_0	-0.1664	0.009104	-18.275	0.000e+00		
	α_1	0.2599	0.013314	19.521	0.000e+00		
	β_1	0.9662	0.004284	225.517	0.000e+00		
TGarch(1,1)	μ	0.09674	0.024489	3.950	7.987e-05	11107.56	11137.53
	α_0	0.10239	0.011654	8.785	0.000e+00		
	α_1	0.06524	0.008523	7.655	2.598e-14		
	β_1	0.83554	0.011056	75.574	0.000e+00		
	γ_1	0.13814	0.015172	9.105	0.000e+00		

Since the test result in Step 1 is positive, we perform Step 2. For simplicity and comparison purpose, we set $p = q = 1$. We employed a GARCH(1,1) model to fit the data. Subsequently, we employed EGARCH(1,1) and TAGARCH(1,1) models. The estimation results are given in Table 4 and it shows that all the coefficients are significant.
In order to assess the performance of the model fit, we perform a residual analysis by taking the standard residual $\frac{\hat{\varepsilon}_t}{\hat{\sigma}_t}$.
If the model is successful at modeling the serial correlation structure in the conditional mean and conditional variance, then there should be no autocorrelation left in the standard residuals and squared standard residuals. The test results are given in Table 5. In all the cases, p-values are greater than *5%* and we conclude that the hypothesis that *"no autocorrelation*" cannot be rejected.

Table 5: Residual Analysis

Residual Analysis: Test for Autocorrelation:-Ljung-Box								
Models	Standardized Residual				Squared Standard Residual			
	Null Hypothesis: No Autocorrelation				Null Hypothesis: No Autocorrelation			
	Stat	P-Value	Dist.	DF	Stat	P-Value	Dist.	DF
GARCH	62.26	8.70e-9	Chi-Sq	12	8.599	0.7368	Chi-Sq	12
EGARCH	62.81	6.89e-9	Chi-Sq	12	10.33	0.5867	Chi-Sq	12
TGARCH	65.67	2.05e-9	Chi-Sq	12	6.047	0.9137	Chi-Sq	12

Since Garch model basically assumes that error ε_t follows a normal distribution and the model is developed mainly to capture the arch effect. Therefore another way of testing the model fit is to check the arch effect of the standardized residual and normality of it. If there is no arch effect then we could conclude that model could capture the arch effects and if there is non-normality then we could conclude that the model is correctly specified. The test results are supplied in Table 6. It indicates that all the models could capture the arch effect of the data but standardized residual does not behave like a standardized normal random variable.

Table 6: Residual Analysis for Arch Effect and Normality

Residual Analysis: Test for Arch Effect and Test for Normality								
Models	Test for Arch Effect				Test for Normality			
	Null Hypothesis: No Arch Effect				Null Hypothesis: Residual is Normal			
	Stat	P-Value	Dist.	DF	Stat	P-Value	Dist.	DF
GARCH	8.329	0.7589	Chi-Sq	12	790.6	0.000	Chi-Sq	2
EGARCH	9.678	0.6442	Chi-Sq	12	762.5	0.000	Chi-Sq	2
TGARCH	6.015	0.9153	Chi-Sq	12	1003	0.000	Chi-Sq	2

Figure 6: Estimated Volatility

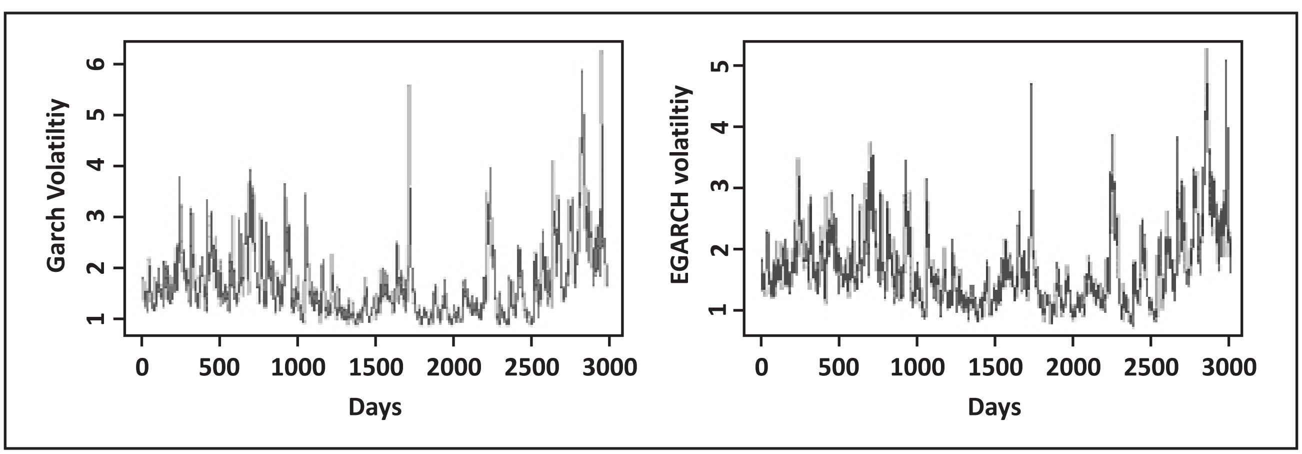
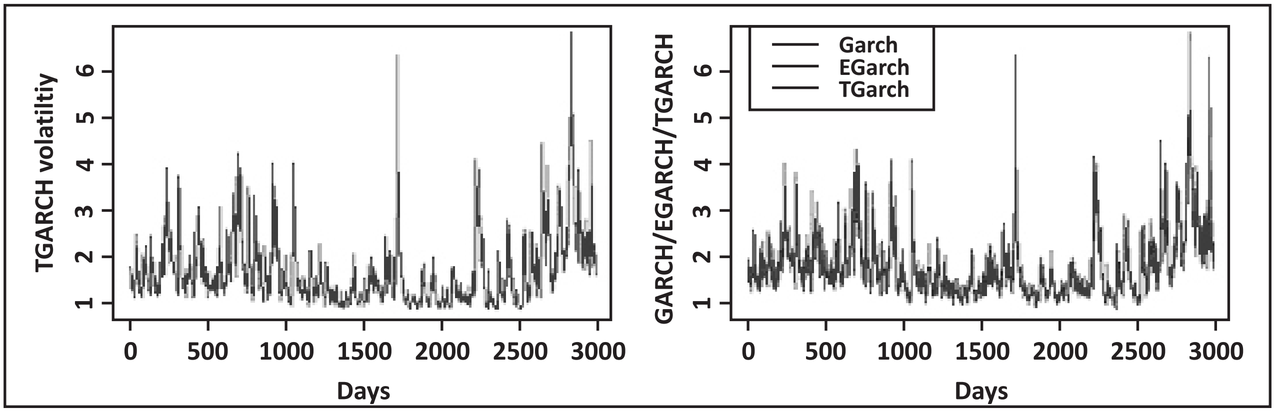


Figure 6 (Contd.): Estimated Volatility



The estimated volatility is shown in Figure 6. The best model among GARCH, EGARCH and TGARCH based on AIC and BIC criterion is TGARCH as it is lowest among all these models.

CONCLUSION

The researcher discussed different mathematical models for modeling volatility in the stock market and application of these models in the Indian context. Our study indicates that GARCH type models could capture the irregular behaviour of the market data and among GARCH models, TGARCH model is more suitable for estimating the volatility in the Indian stock market, especially in the SENSEX market.

ACKNOWLEDGMENT

The author would like to thank his former colleagues in Goldman Sachs for valuable discussions during his stay in the company.

BIBLIOGRAPHY

1. Bollerslev, T.(1986), Generalized auto regressive conditional hetroskedasticity, *Journal of Econometrics*, 31(3), 307-327.

2. Bollerslev, T., Engle, R.F. and Nelson, D.B.(1994), ARCH model, In R.F.Engle and D.C. Macfadden(eds.), *Handbook of Econometrics IV*, 2959-3038, Elsevier Science, Amsterdam.

3. Engle, R.F.(1982), Auto regressive conditional hetroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987-1007.

4. Higgins, M.L and Bera, A.(1992), A class of non-linear ARCH models, *International Economic Review,* 33, 137-158.

5. Glosten, L.R., Jaganathan, R. and Rukle, D.E(1993), On the relation between the expected value and volatility of the nominal excess return on stocks, *Journal of Finance*, 48, 1779-1801.

6. Kumar, S.S.S.(2006), Comparative performance of volatility forecasting models in Indian markets, *Decisions*, 33(2), July-December, 26-39.

7. Nelson, D.B.(1991), Conditional Heteroskedasticity in Asset Returns: a New Approach, *Econometrica*, 59(2), 347-370.

8. Zakonian, J.M(1994), Threshold Hetroskedastic Model, *Journal of Economics Dynamics and Control*, 18, 931 -955.