FORECASTING VOLATILITY IN TEHRAN STOCK EXCHANGE (TSE)

Navid NASR1,*, Mehdi RAZI1

1Industrial Engineering Department, industrial engineering & management system Faculty, Amirkabir University of Technology, Tehran, Iran

Received: 22.03.2015; Accepted: 29.05.2015

Abstract. Within two past two decades, volatility has become a focal point to academics and decision makers. Since volatility was criterion for assessing the risk, it was firstly used by many investors and capital markets traders. Throughout the time as volatility affected the economy and stability of capital markets and as a result of that bonds and foreign exchange markets became important. In this article 9 modeled for forecasting volatility of TEPIX index using Tehran Stock Exchange daily data. Models employed in this article include both simple models such as historical average and complicated models such as ARCH and GARCH groups. 4 measuring tools are used to assess the accuracy of forecasting results of these models. Based on the results of this study, forecasts of GARCH (2,3) presented the most accurate results and GARCH (1,1) and regression models showed less accuracy respectively, while Exponential smoothing and random walk had the worst performance.

Keywords: Forecasting, volatility, tepix, arch, garch

INTRODUCTION

The volatility is related to the rate of price variations with a certain time interval, which has been highly noticed by many researchers, experts, and actives in capital market during recent years but nevertheless this question has no a single and unique answer that which method may purpose the best response to this question. It is tried in this essay to execute the existing and current models for forecasting volatility based on data from Tehran Stock Exchange Index (TEPIX) in order to evaluate them in addition to compare the output from this model versus the real results and to compare the accuracy of their forecasting and at the same time to select the existing optimal model to forecast volatility of this index in order to employ it in other studies in which it requires determining rate of volatility.

The present investigation is intended to acquire appropriate approximation to forecast rate of volatility in Iran Stock Exchange Index; however, at present there is no possibility for transaction of this index in Iranian Stock Exchange Market for religious reasons and it seems improbable as well that during forthcoming years such transaction to be done practically, but nevertheless acquiring such forecasting may be addressed from two aspects: at first place, for those ones who try to create a portfolio similar to the given market (and typically like market index) by diversification of their portfolio in this case knowing rate of volatility may highly contribute to them and on the other hand for many contractual symbols (especially in short term) the mobility of their price is not affected by market index to great extent but severe fluctuations in market affect identically on all symbols (including index) and taking appropriate forecasting may extremely contribute to perceiving risk of market. From empirical perspective, the rate of volatility may play an essential role in pricing of the derived bonds and stocks. According to Black-Scholes Pricing Model, the European call option is a function of volatility. Hence, the

*Corresponding author. Email: n.nasr@aut.ac.ir

Special Issue: Technological Advances of Engineering Sciences

http://dergi.cumhuriyet.edu.tr/ojs/index.php/fenbilimleri ©2015 Faculty of Science, Cumhuriyet University
option markets (which their establishment is included in plans of TSE current year) may be assumed as a place where the contractors transact the volatility. Moreover, in many new models the volatility is also used as one of the main parameters for pricing (particularly for option contracts). Similarly, the iterative series and forecasting of more accurate confidence interval can be achieved by modeling volatility of stock return (ROA).

As it will be explained in the following and in chapter II in this article with more details, various models have been developed and described according to several data from different markets in literature of this subject. It will be dealt with study and comparison the performance and accuracy of forecasting in 9 models of volatility forecasting at Iran Stock Exchange (ISE) and for main index in this market in this survey.

The review of literature subject

The non-parametric techniques are some of the simplest methods, which can be employed to forecasting volatility but due to their weak performance they have not been highly welcomed. In their essays, Pegan and Shirt (2), Kenneth and Dungchul (3) have discussed about way of execution and building the model of forecasting. It may be implied that presentation of ARCH (Autoregressive Conditional Heteroskedasdtasticity) and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) was the start point for continuance and paying attention to forecasting the volatility ratio. ARCH model was suggested for the first time by Engle (4) and then it was developed and completed later by some researchers like Bera and Higgins (5), Bollerslev (6), and Diebold and Lopez (7).

GARCH model was also purposed by Bollerslev and Taylor (8) and it has been developed for several times until now out of which one can refer to EGARCH (Exponential GARCH) model that was introduced by Nelson (9) Of other family of GARCH model, which has been welcomed and utilized, one could imply TGARCH (Threshold GARCH) (Gilson, Jonathan, and Rankel) and or QGARCH (Quadratic GARCH). Akgiray (10) discussed about properties of statistics in stock price and also compared ARCH and GARCH models versus EWMA (Exponential Weighted Moving Average) and historical average model and based on data from American markets he achieved better results according to GARCH model. According to the study that had been done by them based on data from UK market, Damson and Marsh (11) concluded that simple model might act better than Exponential Smoothing and regressive models (It should be noted this point that they had not used models of ARCH family in their investigations). While, Tese (12) with analysis on data from Japanese market and Tese and Tung (13) by using data from Singapore market came to the result that EWMA (Exponential Weighted Moving Average) might purpose more accurate forecasting than GARCH model. West and Cho (14) in an exploration concluded that whether in short term or in long run GARCH model might give better response. Francis and Vandyke (15) examined and evaluated the forecasting potential of GARCH model versus random walk model for 5 European stock markets and found that random walk model might act better (this outcome was repeated by considering crisis and failure in 1987 as well). The interesting point is that according to investigation done by Stach and Watson (16) and based on American macroeconomic data, the random walk model indicated the best performance among all models.

Overall, it can be mentioned that there is a relatively perfect literature about forecasting volatility and many researchers have devoted their investigation to this subject over two past decades. Thus, many econometric models have been examined and utilized in this regard but at the same time, no one could say that all of researchers agreed unanimously in preference of this certain model in
this literature and doubtlessly one could imply that no single model is completely preferred to others.

Concerning to calculation of volatility it should be noted that there are various techniques for computation of monthly volatility (fluctuation). The first suggested solution was a strategy that was introduced by Merton (17) and Perry (18) and accordingly the volatility is obtained as sum of square of daily ratios:

\[ \sigma_T^2 = \sum_{t=1}^{N_T} r_t^2 \]  

(1)

Akgiray (15) offered another solution based on which volatility is calculated as follows:

\[ \sigma_T^2 = \sum_{t=1}^{N_T} (r_t - \bar{r}) \left[ 1 + 0.1 \sum_{j=1}^{N_T-1} \varphi^j \right] \]  

(2)

In this formula, \( r \) denotes the average and \( \varphi \) is the first delay (lag) in correlation. The fact is that the same rule is governed over both equations and both of them emphasize on application of second power of ratio of daily return. Ding et al (19) employed other solution in which the absolute value of daily return ratio was used. Other solution that is concerned with difference among the highest and lowest daily price was suggested by Parkinson (20). Third solution is appropriate because of producing a series that comprises of properties of longer term memory but whereas the long term memories have not been highly noticed thus the given technique has not been widely used. The given response in the second solution is close to the answer in the first model to the great extent. Hence, we will act based on the first solution in this article.

**Description of the utilized model**

In this section, initially the employed models will be explained and in the following the factors and criteria are described for evaluation of performance in these models at Iran Stock and Exchange Market along with estimation the rate of accuracy in their forecasting.

**The used model**

In this part, we will explain about the utilized models and for the sake of more fluent and plain expression, each and every of the used models have been explained separately.

**Random walk**

This model is a type of random walk models as a simple model that is interpreted as follows. According to this model, it is assumed that the last volatility and the existing volatility are the best examples to forecasting of volatility at next period.

\[ \sigma_{T+1}^2 = \sigma_T^2 \quad T = 121, \ldots, 162 \]  

(3)

**Historical average**

If we assume that the expected average is a fixed value, the most optimal forecasting for the future volatility will be as follows:

\[ \sigma_{T+1}^2 = \frac{1}{T} \sum_{T=1}^{T} \sigma_T^2 \quad T = 121, \ldots, 162 \]  

(4)
Previously, this model was the most widely used model but at the same time the recent studies indicate that the expected value may extremely vary over the time (Bollerslev et al, 1997) and this may challenge the validity of this model.

Moving Average

According to historical average model, all of previous observations have the same weight. In moving average model, the closer observations acquire greater weight. In this essay, two types of moving average will be employed i.e. 3-year and 6-year models. The 3-year model is defined as the follows. The moving average model for 6 years period is defined as above but by using data from 72 months period.

\[
\frac{1}{36} \sum_{j=1}^{36} \sigma_{T+1-j}^2 = \sigma_{T+1}^2 \quad T = 121, \ldots, 162 \quad (5)
\]

Simple regression

According to simple linear regression model, the forecasting period (T+1) is defined according to volatility of T-period and as follows where \( \beta_i \) denotes the fixed coefficients based on which the historical data may be obtained:

\[
\sigma_{T+1}^2 = \beta_1 + \beta_2 \sigma_T^2 \quad T = 121, \ldots, 162 \quad (6)
\]

There are two other ways to estimate parameters. In the first method, size of sample remains fixed at 120 when new data is added and thus the last data is deleted. In the second technique, all the existing data are used therefore sample size constantly become greater. The results of both models are extremely close to each other and consequently only the results of model with fixed sample size are reported in this paper.

Exponential smoothing

Exponential smoothing is a simple adaptive technique of forecasting. Unlike regressive method in which the fixed coefficients are employed, the forecasting in exponential smoothing technique adjusts the coefficients based on the previous forecasting errors. The exponential smoothing is expressed as follows:

\[
\sigma_{T+1}^2 = \alpha \sigma_T^2 + (1-\alpha)\sigma_T^2 \quad T = 121, \ldots, 162 \quad (7)
\]

We reach to the following formula by writing the iterative expression:

\[
\sigma_{T+1}^2 = \alpha \sum_{i=1}^{T} (1-\alpha)^i \sigma_{T-i-1}^2 \quad T = 121, \ldots, 162 \quad (8)
\]

Forecasting \( \sigma_{T+1}^2 \) is the weighted average of former value \( \sigma_{T+1}^2 \) in which the weights are reduced exponentially over the time. In this article, \( \alpha \)-value has been calculated in such a way that Root-Mean Square Error (RMSE) to be minimized.

Exponential Moving Average (EMA)

This method is composed of two above-said techniques and according to it, forecasting is acquired according to the following formula:

\[
\delta_{T+1}^2 = (1-\alpha)\delta_T^2 + \frac{\alpha}{L} \sum_{j=1}^{L} \delta_{T+1-j}^2 \quad T = 1, \ldots, 126 \quad (9)
\]

In this article, the values of M=36 and M=72 have been devoted to 3-year and 6-year moving averages, respectively. The \( \alpha \)-value is derived according to minimization of Root-Mean Square Error (RMSE) in forecasting sample.
Weiner Process Model

In this model, it is assumed that the volatility follows up Weiner process. Of presuppositions in such a model, one can refer to Markovian property in such a way that the forecasting value for the next period only depends on present time and the values of former data as well as way of acquiring the current value are not important and this may reduce the number of calculations to the great extent and at the same time lack dependency of on all historical data makes the application of this model easier. This model is expressed as follows:

\[ \Delta V = \mu \Delta t V + \sigma V \varepsilon \sqrt{\Delta t} \] (10)

In this model, \( V \) denotes volatility and \( \Delta t \) is time variation in respective of the previous period; \( \mu \) as average rate of growth and variations in forecasting and \( \sigma \) stands for its volatility. The \( \varepsilon \) is random variable in this model where it follows up standard normal distribution.

ARCH model

ARCH (m) Model that was introduced by Engle (1982) can be defined as follows:

\[ \sigma_n^2 = \gamma V_L + \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \] (11)

Where \( V_L \) denotes variance long-term average and also coefficients \( U \) and \( V_L \) are allocated in such a way that sum of them sets as unit (1). Selection of \( m \) is an essential and important question for implementation task. In this article, \( m \) is selected according to Bayesian Information Criterion (BIC). Like regressive model, the sample size is kept as fixed for this model.

GARCH Model

In most of models which deal with implementation based on ARCH model, m-value should be selected as high and this might increase quantity of calculations so that Bollerslev (1986) introduced GARCH (p, q) Model as follows:

\[ \sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2 \] (12)

In this model, determination of \( p \) and \( q \) is considered as an important issue for implementation. According to what it mentioned in subject of this literature, GARCH model (1,1) is most popular and widely applied model from GARCH family and in order to be able to compare the results of using this model with the results from other studied markets in this literature, this model has been utilized. On the other hand, in order to examine better the effect of change in parameters of this model in this essay, GARCH model (3,2) has been employed to forecasting the volatility of Iranian Stock Market Index as well.

Evaluation criteria

Four criteria are used to evaluate accuracy of forecasting in this essay including:

Root-Mean-Square-Error (RMSE)
\[
RMSE = \sqrt{\frac{1}{I} \sum_{i=1}^{I} (\hat{\sigma}_i^2 - \sigma_i^2)} \quad (13)
\]

Mean Absolute Error (MAE)
\[
MAE = \frac{1}{I} \sum_{i=1}^{I} |\hat{\sigma}_i^2 - \sigma_i^2| \quad (14)
\]

Theil U-statistic
\[
Theil - U = \frac{\sum_{i=1}^{I} (\hat{\sigma}_i^2 - \sigma_i^2)^2}{\sum_{i=1}^{I} (\hat{\sigma}_{i-1}^2 - \sigma_{i-1}^2)^2} \quad (15)
\]

Linex Loss Function
\[
LINEX = \frac{1}{I} \sum_{i=1}^{I} [\exp(-\alpha(\hat{\sigma}_i^2 - \sigma_i^2)) + \alpha(\hat{\sigma}_i^2 - \sigma_i^2)] - 1] \quad (16)
\]

MAE and EMSE are the most popular tools for evaluation of accuracy in forecasting and they are widely used. One of the most major reasons for their popularity is that both ratios are fixed and without change compared to variations and transformations rather than simply calculations. In Theil-U statistic, the forecasting error is standardized by forecasting error of the random walk and value of this statistic becomes 1 for random walk model. But it should be noticed that whereas this is only a type of standardization technique, if value of Theil-U statistic approaches to 1, it is not necessarily a sign of adverse forecasting performance. Most of researchers (e.g. Armstrong and Fields) have analyzed Theil-U statistic and expressed the close formulae for evaluation of accuracy in various forecasting models. However, this statistic does not depend on volume and size of problem, any way it is symmetrical and this may negatively affect on its performance.

Of the paramount reasons for importance and appropriate performance of Linex Loss Function is that the positive differences (errors) are weighted regardless of negative differences (errors) and as a result if \( \alpha > 0 \), then \( \hat{\sigma}_i^2 - \sigma_i^2 > 0 \) is weighted linearly otherwise exponentially. Thus, negative errors are weighted more heavily. In the case of \( \alpha < 0 \), the positive errors are weighted heavier. For more perfect analysis in this essay, values of -20, -10, 10, and 20 have been considered for \( \alpha \); naturally, values of -10 and -20 overestimate this variable while values 10 and 20 may underestimate it.

Implementation of model

In this section, forecasting of the main index in this market is executed by the aid of data of daily index from Iranian Stock and Exchange Market. There are several indices in Tehran Stock and Exchange (TSE) Market out of which TEPIX index is the most perfect as well as the oldest one. TEPIX is an index that measures the related data to the values of negotiable symbols in the bourse market with its corresponding value from base year at any time and expresses the status of Iranian Stock and Exchange Market compared to the base year. Its relevant data exist for almost 14 years so that in this article data from 10 first years (i.e. December 1997 through November 2007) have been employed for estimation of parameters and as samples and thereby data from the last 3.5 years (i.e. December 2007-April 2011) were utilized for estimation and forecasting of volatility.
In Fig (1), diagram of variance was drawn for rate of volatility of this index during the given 10 years.

As it referred in above as well in this essay, 9 different models have been utilized given each of the models of moving average, EWMA, and GARCH were executed in both forms then totally 11 models have been implemented. It should be mentioned that regarding Exponential Smoothing and EWMA models calculation of $\alpha$-parameter has been adjusted in such a way that Root Mean Square Error (RNMSE) to be minimized. This trend has been employed for determination of $\alpha$, $\beta$, and $\gamma$ in ARCH and GARCH models. Similarly, to determine parameters in ARCH model (m-value), BIC analysis was utilized where finally the value ($m = 10$) was selected for ARCH model.

As it also referred to this point as above, some differences among the present essay from what it has been yet mentioned in this literature at first place is related to its application for Iranian Capital Market where its data has been rarely used in this literature and at the same time one can refer to implementation and comprehensive comparison among the given models in this literature (ranged from simple models like random walk to complex models e.g. stochastic volatility). The evaluation indices are some of other cases, which could be mentioned for application that they have been less addressed in this literature including application of Linex loss function that is asymmetric index with higher flexibility in evaluation of accuracy in forecasting cases. The summary of the given results derived from these models on data is given in Tables 1 and 2. RSME criterion indicates that GARC (3,2) Model may forecast the best answer and then GARCH (1,1) Model is ranked after this model. As it also referred to this point in section of review of literature, despite of its very simple calculations, random walk model may give appropriate answers but these answers are not suitable model for TEPIX index and they are not close to real results. The regressive model has the best performance according to MAE statistic and then ARCH model is placed. It is interesting point that according to this statistic, Exponential Smoothing model is ranked at last position. The other interesting point is that this statistic is the product of appropriate performance of Weiner Process Model that is ranked at fourth place. In Theil-U statistic, only Exponential Smoothing Model acts worse than random walk model and its value became greater than unit (1). In this statistic, GARCH group has also the best performance. But unlike the former criterion, ARCH model has no appropriate performance.
Table 1. Table of Results and comparison between results from implementation of models based on Theil-U, MAE, RSME criteria.

<table>
<thead>
<tr>
<th>Model</th>
<th>RSME Value</th>
<th>Rank</th>
<th>MAE Value</th>
<th>Rank</th>
<th>Theil-U Value</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random walk</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical average</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moving average (M-33)</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moving average (M-73)</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exponential smoothing</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EWSMA (M-36)</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EWSMA (M-73)</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter Process</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH (1,1)</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH (2,3)</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Table 2, the models have been examined and compared according to Linex statistic that is an asymmetric criterion. The most marvelous point is that the totality of ranking the model is identical and generally GARCH models group has the better performance while Exponential Smoothing and random walk models possess the worst performance. As it also mentioned above, the underestimate fine is considered for α - positive value in Linex loss function and vice versa, but according to data from Iranian Stock Market, these models are too robust (sensitive) to variation of α-value so the final answer does not highly vary and this means that in this case the symmetric and asymmetric criteria do not differ particularly and in fact dissymmetry of Linex criterion does not create high value-added.

Table 2. Table of Results and comparison between results from implementation of models based on Linex statistic.

<table>
<thead>
<tr>
<th>Model</th>
<th>Linex 1</th>
<th>Linex 2</th>
<th>Linex 3</th>
<th>Linex 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSME Value</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
</tr>
<tr>
<td>MAE Value</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
</tr>
<tr>
<td>Theil-U Value</td>
<td>0.00210209</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
</tr>
<tr>
<td>Linex 1 Rank</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
</tr>
<tr>
<td>Linex 2 Rank</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
</tr>
<tr>
<td>Linex 3 Rank</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
</tr>
<tr>
<td>Linex 4 Rank</td>
<td>11.99815557</td>
<td>13.99781742</td>
<td>15.99795534</td>
<td></td>
</tr>
</tbody>
</table>

CONCLUSION

Discussion

During recent years, forecasting volatility is assumed as one of the foremost fields of researches in financial markets and due to a great number of applications of volatility in various models, their forecasting was employed for various applications. The present article has examined and tested 9 different models and employed these models for data during 10-years period in main index (TEPIX) at Iranian Stock and Exchange Market and to forecast 3.5 years future period in this market. According to the acquired results from forecasting cases, GARCH (2,3) model achieved the first rank based on RMSE, Theil-U, and Linex criteria (in all four modes) and then GARCH (1,1) and regressive models are placed. This point should be also taken into consideration that these models are difficult whether theoretically and in terms of computations therefore the advantages of a model like random walk model should not be forgotten in which despite of a few calculations, it gives relatively appropriate answer while Exponential Smoothing and random walk models have the inappropriate performance and they have forecast the most improbable response.
Suggestions

With respect to the aforesaid topics in above, it seems that the following cases could be appropriate fields for the future studies. Given welcoming to random processes models, it seems that review and evaluation of other forecasting models such as Stochastic Volatility and their comparison with the existing techniques may provide the favorable ground for the studies. Making effort to create appropriate methods for estimation of optimal parameters in ARCH and GARCH models may affect on improving performance in these methods and their further application. Likewise, employing other existing criteria for evaluation and study on accuracy and precision of forecasting for better and more accurate assessment of forecasting models may contribute to better comparison among these models.

REFERENCES