Application of the generalized autoregressive conditional heteroskedasticity model in forecasting the stock price index: Case study from Vietnam

Le Thi Minh Huong¹,², Atta Ullah³, Muhammad Asif Khan³, Do Thi Thao⁴

¹. School of Economics, Huazhong University of Science and Technology, China
². Department of Economics, Danang Architecture of University, Vietnam
³. School of Management, Huazhong University of Science and Technology, China
⁴. School of Hospitality & Tourism, Hue University, Vietnam

Corresponding Author: Le Thi Minh Huong

Abstract: Aims and Objectives: The purpose of this study is to forecast the Vietnam stock price (VN index).

Methodology: Authors used the generalized autoregressive conditional heteroskedasticity (GARCH) model and utilized inspection standards to find the best model.

Results: The study gives the results that the GARCH (1,1) and the Threshold GARCH (TGARCH) (1, 1) models are appropriate to carry out the forecast.

Conclusion: The results also showed that the Vietnam stock market had leverage impact on asymmetric information. The model at mean value GARCH-M indicates that the mean value of the VN index is not affected by its variance, in other words, independent on the risk of price volatility from the market.

Keywords: GARCH, forecasting, Vietnam stock market, VN index, volatility.

I. Introduction

The prediction problem starts to appear from ancient times and is growing actively until now, becoming an indispensable part for human activities in the context of information explosion. All areas of life around us need forecasts, such as in the field of meteorology and hydrology, weather temperature forecasting help a lot for the economy as well as avoiding losses and great harm caused by nature. In the financial sector, if we predict that the rise or fall of the exchange rate, the currencies or the price of a stock will bring many benefits. In short, forecasting provides the necessary bases for planning; if there is no predictive science, the future that people plan will not be highly persuasive.

Vietnam stock market was established in July 2000; the market has developed actively, increasingly improved in structure, expanded in scale, become an important medium and long-term capital channel in the economy. By the end of June 2018, the stock market capitalization reached 3,889 million VND, equivalent to 77.7% of GDP. After ten years of the global economic crisis, April 10, 2018, for the first time in trading history, the VN index peaked at 1211 points. Previously, this index re-established the highest level before the financial crisis in 2007 reached 1171 points. However, there were also times when the market plunged to 235 points on February 24, 2009. 2018 is the year when Vietnam's economy grew the most in the last ten years. However, the stock market this year is contrary to the general trend of the domestic economy. This stems from the instability from the international market such as escalating trade war; the US Federal Reserve raised interest rates causing capital flows to rotate, the US dollar rose, to the high instability. The commodity market, especially crude oil, has reversed all previous forecasts. Investors and policymakers have always questioned how the stock market price in Vietnam will fluctuate? Which model should we use from past data, as well as psychological preparation for the risks they face in the future?

In this case, the forecast for stock prices is useful for investment purposes in Vietnam. However, the characteristics of Vietnamese stock prices fluctuate considerably by changing events from home and abroad. Volatility is the situation in which the variance of conditions changes between a very high and low-value state, in theory, when dealing with time series, it is vital to use it to predict volatility or variance changes over time. The autoregressive conditional heteroskedasticity (ARCH) model is an economic model describing the sequence with the variance that changes over time (Engle, 1982). The family of GARCH model systems was developed to capture group dynamics or fluctuations and predict future volatility (Bollerslev, 1986).
We organize this paper as follows; section 2 summarizes the research overview, section 3 designates the data and methods for research. Experimental outcomes are described in section 4. Lastly, section 5 gives some inferences.

II. Material And Methods

1. Material

There are many studies on fluctuations in the stock market implemented in the world and Vietnam. The study of changes in the VN index in Vietnam stock market plays a vital role for fund managers and investors in the market. Looking back at Vietnam’s stock market in the past 20 years, there have been many changes. When studying stock price volatility, we apply the ARIMA model to continue analyzing the ARCH model, thereby forecasting risks through several GARCH models.

The research that has collected data is the price index of trading closed the VN-Index by the day from January 2, 2009, to Dec 28, 2018, includes 2494 observations that were used to measure the change of the VN index daily. The second period, within more than one week (from December 31, 2018, to January 8, 2019), is for out-of-sample forecasts.

Figure 1: Fluctuation of VN index from 2009 - 2018

Looking at Figure 1 the graph of stock prices at Vietnam (2009 – 2018) has an increasing trend. Visually, based on the above graph, it can be deduced that the condition of expected value does not depend on the time that has been violated, and this is a time series without stationary. To be able to perform forecasting operations, we need to convert the original series by calculating the return of the stock price index. The return rate $R$ of the VNI stock price index may be a stationary sequence and may have an impact on the ARCH. The formula for calculating the rate of return is as follows: $RVNI = \log (VNI / VNI (-1))$.

2. Procedure methodology

ARCH model

The ARCH model (Engle, 1982) suggested that it is best to simultaneously model the mean and variance of the data series when suspecting that variance values change over time. The ARCH (1) model will be written as:

\[ Y_t = \beta_1 + \beta_2 X_t + u_t \]

\[ u_t \sim N(0, h_t) \]

\[ h_t = \gamma_0 + \gamma_1 u_t^{2-1} \tag{2} \]

Here, equation (1) is for estimating the mean (e.g., stock price index) and equation (2) is for calculating the variance value (e.g., stock market risk). Note, to simplify the expression of the comparison of the variance equation, from here on we use the symbol $ht$ instead of $\sigma_t^2$.

The ARCH model (1) assumes that if there is a big shock occurring at time $t-1$, the $u_t$ value (absolute value or square) will also be more significant. That is, when $u_{t-1}^2$ is big/small, $u_t$ variance will be big/small. Regression coefficient $\gamma_1$ is nonnegative because the variance is always positive. Conditional variance can depend on many previous lags since each case can create a different ARCH process. And the general case will be ARCH (q) expressed as follows:

\[ h_t = \gamma_0 + \gamma_1 u_{t-1}^2 + \gamma_2 u_{t-2}^2 + \cdots + \gamma_q u_{t-q}^2 = \gamma_0 + \sum_{j=1}^{q} \gamma_j u_{t-j}^2 \tag{3} \]

The estimated coefficients $\gamma_j$ must be positive because the variance is always positive.
Model GARCH

According to (Engle, Kroner, & Engle, 1995), the ARCH model is limited when it seems to be more like the moving average model than the Autoregressive integrated model. (Fan and Yao, 2003) said that Arch (p) is only suitable for financial models with large enough ligand this ensures extensions for the Arch model. Therefore, a new idea that they should include variables with the t-lag of the conditional variance in the variance equation (Bollerslev, 1986). Predictors tend to choose more popular GARCH models for the application both of short and long time series. Model GARCH (p, q) can be written as:

\[ Y_t = \beta_1 + \beta_2 X_t + u_t \]
\[ u_t \sim N(0, h_t) \]
\[ h_t = \gamma_0 + \sum_{i=1}^{p} \delta_i h_{t-i} + \sum_{j=1}^{q} \gamma_j u_{t-j}^2 \]  

Where (4) is the mean equation, (5) is the variance equation, this equation shows that the variance h_t now depends on both the past value of shocks, the representation by the hysteresis variables of the squared noise class, and the self-past values of h_t, represented by lt-i variables. If p = 0, meaning that the order of AR is 0, then the model GARCH (0, q) is simply an ARCH (q) model. The simplest form of the GARCH model (p, q) is

\[ \text{GARCH (1, 1) model of the form:} \]
\[ h_t = \gamma_0 + \delta_1 h_{t-1} + \gamma_1 u_{t-1}^2 \]

To ensure that the variance has a non-negative condition, the coefficients \( \delta_1, \gamma_1 \) must be \( \geq 0 \). With, \( \delta_1 \) measures the variation that can occur in the next period. If \( \delta_1 \) is high, it can show that changes in the market change are significant. If the coefficient \( \gamma_1 \) is high, it indicates that there is a long-term fluctuation.

Besides, if \( \delta_1 \) is high and \( \gamma_1 \) is low, the fluctuation is extreme. If the sum of \( \delta_1 \) and \( \gamma_1 \) approaches 1, a shock at time t will exist for a long time in the future. If the sum of \( \delta_1 \) and \( \gamma_1 \) is less than 1, it will lead to frequent changes in the long run. If the sum of \( \delta_1 \) and \( \gamma_1 \) equals 1, shocks will temporarily change future values. The GARCH model relies on the dependency of the series of changes, giving future observations based on past observations, so the GARCH model based on variance changes over time.

GARCH-M model

GARCH model at mean values allows conditional averages to depend on their conditional variance. For example, consider the behaviour of investors in the form of “risk”, and therefore, they tend to require an additional risk compensation fee as part of the compensation to decide to hold a risky asset. Thus, risk compensation is a variable function with risk; that is, the higher the risk, the greater the risk compensation cost. If the risk is measured by the oscillation level or by the conditional variance, then the model GARCH (0, q) is simply an ARCH (q) model. The simplest form of the GARCH model (p, q) is

\[ \text{GARCH-M model (p, q) can be given by:} \]
\[ Y_t = \beta_1 + \beta_2 X_t + \theta h_t + u_t \]
\[ u_t \sim N(0, h_t) \]
\[ h_t = \gamma_0 + \sum_{i=1}^{p} \delta_i h_{t-i} + \sum_{j=1}^{q} \gamma_j u_{t-j}^2 \]  

Another form of the GARCH-M (p, q) model is that, instead of using the series of variance in the average equation, we use the standard deviation of the conditional variance sequence as follows:

\[ Y_t = \beta_1 + \beta_2 X_t + \theta \sqrt{h_t} + u_t \]
\[ u_t \sim N(0, h_t) \]
\[ h_t = \gamma_0 + \sum_{i=1}^{p} \delta_i h_{t-i} + \sum_{j=1}^{q} \gamma_j u_{t-j}^2 \]

T-GARCH model

The T-GARCH model is developed by (Jean-Michel Zakoian, 1994) and (Lawrence R. Glosten, Ravi Jagannathan, and David E. Runkle, 1993). This model considers the asymmetry between positive and negative shocks. And this is also a way to test the effectiveness of the market. To do so, these scholars propose to include the equation of the variance of a dummy variable interacting between the squared noise class and the dt variable dt, where dt is equal to 1 if ut < 0, and 0 if ut > 0. If the coefficient of this interaction variable is statistically significant, there will be differences in different shocks. From this idea, the variance equation in the T GARCH model (1,1) will look like this:

\[ h_t = \gamma_0 + \delta_1 h_{t-1} + \gamma_1 u_{t-1}^2 + \theta_1 u_{t-1}^2 d_{t-1} \]  

If the coefficient \( \delta_1 \) is statistically significant, good news and bad news will have different effects on the variance. Specifically, good news only affects \( \Delta \), while bad news has an impact (\( \Delta \) + \( \Delta \)). If \( \delta_1 > 0 \), then we can say that there is an asymmetry in the impact between good news and bad news. Conversely, if \( \delta_1 = 0 \), then the impact of news is proportional. High-level TARCH can be expressed as follows:

\[ h_t = \gamma_0 + \sum_{i=1}^{p} \delta_i h_{t-i} + \sum_{j=1}^{q} (\gamma_j d_j) u_{t-j}^2 \]  

DOI: 10.9790/5933-1005050110 www.iosrjournals.org 3 | Page
Application of the generalized autoregressive conditional heteroskedasticity model in forecasting ...

Forecast of variance

Static forecasting can predict for the next period \( n + 1 \); dynamic forecasting has the advantage of forecasting for a longer period of out-of-sample. For simplicity, the predictive GARCH (1,1) model is defined as:

\[
h_{t+1} = \gamma_0 + \delta h_t + \gamma_1 u_t^2
\]  

(13)

III. Results and Discussion

1. Stationary of data and ARIMA model

Table 1: Testing the stationary of VNI data series

<table>
<thead>
<tr>
<th>Unit root test of RVNI</th>
<th>Phillips-Perron</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller</td>
<td>Phillips-Perron</td>
<td></td>
</tr>
<tr>
<td>t-Statistic</td>
<td>Prob.</td>
<td>Adj. t-Stat</td>
</tr>
<tr>
<td>-42.89961</td>
<td>0.0000</td>
<td>-43.30198</td>
</tr>
</tbody>
</table>

With results from Table 1 has been said that: The data series is stationary at level I (0) with three significant 1%, 5%, and 10%. Therefore, we can rely on this series to select the appropriate ARIMA model. There are many methods to determine the parameters (p, q) in the model ARIMA. Among them, a commonly used method is based on the degree the lag of ACF and PACF on the correlation diagram.

From Figure 2, using the ACF autocorrelation coefficients to select the order q for MA, the autocorrelation coefficient, PACF alone to choose the p-level for AR. RVNI is stationary, so we take p and q as those. The value is outside the confidence interval (5% significance level), so consider the ARMA (1, 1), ARMA (0, 1), ARMA (1, 0) models based on the test criteria AI, SC and HQC.

Table 2: Select the appropriate ARIMA model

<table>
<thead>
<tr>
<th></th>
<th>ARMA (1, 1)</th>
<th>ARMA (0, 1)</th>
<th>ARIMA (1,0)</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>R^2</td>
<td>0.022790</td>
<td>0.021292</td>
<td>0.022476</td>
<td>Chose ARMA (1, 0)</td>
</tr>
<tr>
<td>AR(1) coefficient</td>
<td>0.274407</td>
<td>0.149882</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA(1) coefficient</td>
<td>-0.127605</td>
<td>0.142021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>-5.8867642</td>
<td>-5.886915</td>
<td>-5.888124</td>
<td></td>
</tr>
</tbody>
</table>
Application of the generalized autoregressive conditional heteroskedasticity model in forecasting...

Conclusions are drawn from Table 2 shows that: The optimal model is the ARIMA (1, 0) model because the critical value is smaller than the ARMA(0, 1) model. The ARMA (1,1) model was excluded because it was not statistically significant. However, the ARIMA model generally has the disadvantage of not being able to overcome the phenomenon of heteroskedasticity of the time series. Therefore, we should check whether the ARMA (1, 0) model is selected to be affected by this effect.

Check the variance of the variation error made by testing Engle’s ARCH. Testing Engle’s ARCH is a check on the right side (if Test Statistics is greater than the allowed threshold - Critical Value rejects the assumption H0). The verification value of Engle’s ARCH audit follows the chi-squared distribution with the degree of freedom by the lags of the test.

With the results obtained, the null hypothesis does not exist. This means there is a phenomenon of conditional Heteroskedasticity. That is, the ARCH effect exists with small p-value (0.0000 <0.05). The value of the squared computation of 186.9278 is too high compared to the Table of Values value at the 1% significance level (6.6349), so we reject the H0 hypothesis. That is, the ARMA model (1, 0) affects the ARCH. Because this time series has an ARCH effect, we will change the way of estimating the Least Squares model to the ARCH method.

2. Choose ARCH and GARCH model
   a. Estimating Arch model

<table>
<thead>
<tr>
<th>Table 3: Estimating an appropriate ARCH model</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R^2 )</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>R²</td>
</tr>
<tr>
<td>AIC</td>
</tr>
<tr>
<td>SC</td>
</tr>
<tr>
<td>Prob. Chi²</td>
</tr>
</tbody>
</table>

Conclusions are: \( \text{ARCH (7)} \) is the best model.

Results from Table 3 come from continuing to increase the number of lags to 2, 3, 4, 5, 6, 7, 8 we found that there is a possibility that the 7th lag is the optimal lag, since the estimated coefficients in the sub-regression model are significant at 1%, and other statistics such as adjusted \( R^2 \), AIC, SC, etc., there is no significant difference from the validity lag smaller than the previous lags. \( \text{ARCH (8)} \) is not statistically significant, so we remove \( \text{ARCH (8)} \) and more. However, using too many lags, not always the optimal solution, so in such cases, we prefer to use the GARCH model (p, q).

<table>
<thead>
<tr>
<th>Table 4: Verification about conformity of the ARCH (7) model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unit root test of residual series</strong></td>
</tr>
<tr>
<td>Augmented Dickey-Fuller</td>
</tr>
<tr>
<td>t-Statistic</td>
</tr>
<tr>
<td>-49.33778</td>
</tr>
</tbody>
</table>

From Table 4, the results illustrate that the residual is stationary at the initial level I (0), which is good to use in subsequent analysis. However, the GARCH model with small lag is a compact representation of the ARCH model (q), with q extending infinitely and coefficients that tend to decrease. For this reason, we should use the GARCH model with small lag instead of higher ARCH models because of the GARCH model with a short lag; we have fewer coefficients to estimate. Therefore, it helps to limit the possibility of losing some degrees of freedom in the model.
b. Estimating the GARCH model

![Volatility of the stock price index return](figure3.png)

Figure 3: Volatility of the stock price index return

From the Figure 3, we can see that the VN index has the characteristics of clustering volatility. With this volatile property, the GARCH model used is very suitable for forecasting. To determine the appropriate GARCH model, we need to choose the lag (p, q) for the (GARCH, GARCH-M, TGARCH) models.

<table>
<thead>
<tr>
<th>Variable</th>
<th>GARCH (1, 1)</th>
<th>GARCH(2,1)</th>
<th>GARCH(3,1)</th>
<th>GARCH(1, 2)</th>
<th>GARCH(1, 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef.</td>
<td>4.29E-06 0</td>
<td>4.28E-06 0</td>
<td>4.19E-06 0</td>
<td>4.31E-06 0</td>
<td>3.86E-06 0</td>
</tr>
<tr>
<td>Prob.</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Coef.</td>
<td>0.138248 0</td>
<td>0.137586 0</td>
<td>0.136117 0</td>
<td>0.137290 0</td>
<td>0.136244 0</td>
</tr>
<tr>
<td>Prob.</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Coef.</td>
<td>0.838882 0</td>
<td>0.845296 0</td>
<td>0.961029 0</td>
<td>0.001351 0</td>
<td>0.019429 0</td>
</tr>
<tr>
<td>Prob.</td>
<td>0.9730</td>
<td>0.9641</td>
<td>0.9641</td>
<td>0.0000</td>
<td>0.6211</td>
</tr>
<tr>
<td>Coef.</td>
<td>-0.005671 0</td>
<td>0.251156 0</td>
<td>0.5000 8</td>
<td>0.838382 0</td>
<td>-0.029128 0</td>
</tr>
<tr>
<td>Prob.</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Coef.</td>
<td>-6.140652</td>
<td>-6.139850</td>
<td>-6.139850</td>
<td>-6.139850</td>
<td>-6.139354</td>
</tr>
<tr>
<td>Prob.</td>
<td>0.131455 0</td>
<td>0.3547</td>
<td>0.3547</td>
<td>0.852713 0</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 5: Results of a selection of GARCH models

Based on Table 5, the study proposes to remove GARCH (3, 1), GARCH (1, 2), GARCH (1, 3), GARCH(2, 1) models because these models provide negative coefficients and p-value higher than 5%. GARCH (1, 1) models have good results. It has resulted in non-zero constants; the values of AR are positive to ensure the condition. Besides, the coefficients $\delta_1$ and $\gamma_1$ are statistically significant (less than 1%), which may indicate that the fluctuation of the previous period may explain the variations of this period. The total of these two coefficients is less than 1, ensuring the stationary of the conditional variance. Also, when checking the ARCH calculation of the GARCH model (1, 1), see a significant p-value more than 0.05, the model does not have heteroscedasticity.
We proceed to compare whether the model is better between ARCH (7) and GARCH (1, 1).

Looking at the graph in Figure 4, we see that it seems like the ARCH (7) model and GARCH (1, 1) model are very similar. However, if we look closely, you will realize that the GARCH (1, 1) model has a relatively smooth estimation variance value than the ARCH (7) model. Thus, it is best to use the GARCH model instead of the higher-level ARCH models as we have just saved the number of degrees of freedom (especially when the number of observations is less) and more convenient in the forecast (reduce calculation).

The mean equation has conditions and the equation of conditional variance of the VN index series estimated from the GARCH model (1, 1) in the following form:

\[
RVN = 0.000530917310726 + 0.126566096427\times RVNI(-1) + Ut
\]

With Ut obey standard distribution rules, Coefficient AR (1) is 0.126566096427 (greater than 0) proving that the VN Index value at present has a positive correlation with the previous index value. Besides the mean value, the GARCH model (1, 1) also estimates both the conditional variance with the equation is

\[
GARCH = 4.29392684568e^{-06} + 0.138247991675\times RESID(-1)^2 + 0.838882158597\times GARCH(-1).
\]

The equation consists of two components, which are past information and conditional variance in the past. The coefficients of these components in the equation are positive (the coefficient ARCH (1) is 0.138247991675, the coefficient GARCH (1) is 0.838882158597) and are statistically significant, indicating that the variance has the current events positively correlated with past information and past conditional variances of the preceding moment. In other words, the current level of VN-Index fluctuation depends on the change of the VN-Index in the past (represented by the ARCH coefficient), both depending on the volatility level of this change in the past (expressed by GARCH coefficient). In which, the coefficient GARCH (1) is greater than the ARCH coefficient (1), indicating the impact of the past conditional variance on the current conditional variance is stronger than the effects of the previous information.

3. Estimating T-GARCH model and GARCH-M model

The most significant disadvantage in GARCH models is that they are assumed to be symmetrical. This means that these models only care about the absolute value of shocks, not their ‘mark’ (because we handled noise/residuals in the squared form). Therefore, in ARCH/GARCH models, a sharp shock with a positive value affects the fluctuation of the data series completely like a severe shock with a negative value. However, experience shows that, in finance, when having adverse shocks (or bad news), investors are often pessimistic, depressed, and waiting to prosper from the market passively. This impacts more strongly and persistently than positive shocks (or good news). Therefore, people try to model the difference in this influence, and we have a T-GARCH model.

With results from Table 6, the regression coefficient of the interaction variable is positive (0.062725), and statistically significant, proving that there is a noteworthy difference between the impact of the news and good news on stock index VN index. We can also say there is a leverage effect on the VN index in Vietnam.

The results of model estimation GARCH-M (1, 1) is not statistically significant. It displays that the coefficient of variance (2.701531) in the mean equation with p-value is higher than 0.05. This may indicate that the GARCH-M (1, 1) model is not appropriate in this case. Therefore, T- GARCH (1, 1), which model is better for predicting about the statistics of the VN index.
Application of the generalized autoregressive conditional heteroskedasticity model in forecasting...

Table 6: Estimating T-GARCH model and GARCH-M

<table>
<thead>
<tr>
<th>Variable</th>
<th>T-GARCH</th>
<th>GARCH-M</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>0.000379</td>
<td>0.000266</td>
<td>- Leverage effect on T-GARCH model</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.128066</td>
<td>0.127447</td>
<td>- The GARCH-M model is not statistically significant</td>
</tr>
<tr>
<td>$\theta$</td>
<td>2.701531</td>
<td></td>
<td>- T-GARCH is the appropriate model selected for forecasting</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>4.85E-06</td>
<td>4.26E-06</td>
<td></td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.832505</td>
<td>0.137783</td>
<td></td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>0.108008</td>
<td>0.839530</td>
<td></td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.062725</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\vartheta_1$</td>
<td>0.062725</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * means that coefficient is no statistically significant

4. Predict the stock price index in the coming time

Table 7: Some standards check the appropriateness of the forecasting models

<table>
<thead>
<tr>
<th>Standards</th>
<th>T – GARCH (1, 1)</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.012873</td>
<td>T-GARCH(1, 1) is good model</td>
</tr>
<tr>
<td>MAE</td>
<td>0.009321</td>
<td></td>
</tr>
<tr>
<td>Bias Proportion</td>
<td>0.000011</td>
<td></td>
</tr>
<tr>
<td>MAPE</td>
<td>180.8877</td>
<td></td>
</tr>
</tbody>
</table>

RMSE and MAE indicators at Table 7 reflect the difference between forecast value and actual value. The closer these indicators are to 0, the more accurate. Here, both RMSE and MAE are all larger than 0, but still low so it is acceptable. MAPE also has the same meaning as the above two criteria but is in percentage form and also located permitted levels. For the coefficient Bias too tiny, only 0.0000011. In general, these indicators show that the model can be used to make forecasts for the time series of the VN-Index.

Table 8: Results of predicting the average and the variance value of the VN index in the next week

<table>
<thead>
<tr>
<th>Date</th>
<th>VNIf</th>
<th>RVNIf</th>
<th>Forecast of Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/28/2018</td>
<td>892.54</td>
<td>-0.009223</td>
<td>0.0001024</td>
</tr>
<tr>
<td>12/31/2018</td>
<td>891.78</td>
<td>-0.0008504</td>
<td>0.0001044</td>
</tr>
<tr>
<td>01/01/2019</td>
<td>891.98</td>
<td>0.0002218</td>
<td>0.0001063</td>
</tr>
<tr>
<td>01/02/2019</td>
<td>892.30</td>
<td>0.0003591</td>
<td>0.0001082</td>
</tr>
<tr>
<td>01/03/2019</td>
<td>892.64</td>
<td>0.0003767</td>
<td>0.0001100</td>
</tr>
<tr>
<td>01/04/2019</td>
<td>892.97</td>
<td>0.000379</td>
<td>0.0001117</td>
</tr>
</tbody>
</table>

Figure 5: Graph of results of the VN-Index forecast

The forecast results from the model at Figure 5 and Table 8 show that in the next week, the stock price index will decrease compared to the last day in the past data sample, but then will increase slightly. The degree of price volatility in the T-GARCH (1, 1) will be less than that of the GARCH (1, 1), but the difference is negligible.

The VN index in early January is forecast to fall from the end of December: nearly 1 point lower than the actual value on December 28. The trend of the market in the week is predicted to increase again on the
weekends. But in general, during the week, the gaining trend is still dominant. However, the level of fluctuation is not too large; about 1 point per day increases.

IV. Conclusion

Our study gives some main findings as follows: First, for predicting stock price indexes in Vietnam, the use of GARCH (1, 1) supposedly suitable for application in the future. The T-GARCH (1, 1) is also considered to be a good model for forecasting stock prices based on other evaluation criteria. Second, the ARMA model proved inappropriate because of the impact of ARCH through the test results of variance. Third, although the ARCH(7) model satisfies the conditions, is an excellent predictive model. However, the number of lags is too large, which makes the use of the model influence the estimation values, due to significantly reduce the total of degrees of freedom of the regressive model. In some cases, the time series is too short; this is more serious, such as predicting the new share price will be inaccurate. Because of this, predators often prefer using a GARCH model that tends. Finally, depending on the judgment of investors as well as the criteria for selecting the reliability that investors will prefer to choose which of the two models of GARCH (1, 1) or T-GARCH (1, 1).

As for the forecast results, in the first week of January, also the first trading week of the stock market of the new year, the market is forecasted to increase. When conducting further research, some methods we can apply to the Vietnam stock market such as machine learning approach, the network based on the adaptive system, or the use of neural networks, and so on.

References

Application of the generalized autoregressive conditional heteroskedasticity model in forecasting the stock price index: Case study from Vietnam

Le Thi Minh Huong


