MODELING VOLATILITY USING GARCH MODEL: NASDAQ-100 APPLICATION

Marina Dobrota¹, Ana Poledica¹, Milica Bulajić¹, Bratislav Petrović¹
¹Faculty of Organizational Sciences, University of Belgrade

Abstract – Modeling and forecasting stock market volatility is very important issue in many financial applications, and it has been the subject of numerous studies in financial markets, both theoretical and practical. In finance, volatility can be interpreted as a measure of fluctuation in a financial security price around its expected value. The main problem, which arises from the very definition of volatility, is the unstable nature of observed time series and its heteroscedasticity, making it impossible to apply certain time series models. First models to overcome these difficulties were ARCH and GARCH, followed by various extensions and modifications of GARCH model. This study empirically investigates the forecasting performance of GARCH model for the NASDAQ-100 stock index over the period Jan 2005–June 2011. NASDAQ-100 stock index reflects companies listed on the NASDAQ stock market across major industry groups including computer hardware and software, telecommunications, retail/wholesale trade and biotechnology. Future research can be directed towards integration of the proposed volatility forecasting model with artificial neural networks in order to capture nonlinear patterns.

1. INTRODUCTION

Volatility is an essential input to many financial decision making models, thus it is has been the subject of many research in financial markets, both practical and theoretical. It is one of the primary inputs to a wide range of financial applications from risk measurement to asset and option pricing [1]. Investment decisions in financial markets strongly depend on the forecast of expected returns and volatilities of the assets. Volatility is the relative rate at which the price of a security moves up and down around its expected value. In the terms of this research, volatility is a statistical measure of the dispersion of returns for a given security or market index [2].

Volatility can be measured by using the standard deviation or variance between returns from that same security or market index. Commonly, the higher the volatility, the riskier the security. This brings us up to the main problem, which arises from the very definition of volatility: the unstable nature of observed time series and its heteroscedasticity. In fact, most models for forecasting (such as AR, MA, ARMA, etc.) are assuming that the variance is a constant, i.e. homoscedasticity of a time series.

The issue of modeling time series with changing variance and heteroscedastic errors was always vague. First models to overcome these difficulties were Autoregressive Conditional Heteroscedasticity (ARCH) introduced by Engle [3]. Still, it has long been recognized that a number of stylized facts, such as persistence, volatility clusters, time-varying volatility, and leptokurtic data behavior, often characterizes returns volatility. The introduction of Generalized ARCH (GARCH) model by Bollerslev [4] has created a new approach for accommodating these temporal dependencies for financial econometricians, becoming a popular tool for volatility modeling and forecasting [5]. However, despite the success of the GARCH model, it has been criticized for failing to capture asymmetric volatility. This limitation is overcome by introducing more flexible volatility treatments by accommodating the asymmetric responses of volatility to positive and negative shocks. This more recent class of asymmetric GARCH models includes the Exponential GARCH (EGARCH) model of Nelson [6]. These models have been extensively used in finance and economics, specifically applicable to modeling financial market data, which are highly volatile [1], or quantifying risks of VaR (value at risk) under stress times [7].

This study primarily focuses on anticipation the volatility of return on NASDAQ-100, using suitable forecasting models. We have chosen the NASDAQ-100, since the index reflects companies across major industry groups including computer hardware and software, telecommunications, retail/wholesale trade and biotechnology [8]. The NASDAQ-100 is a stock market index based on market capitalization. It includes 100 of the largest non-financial companies listed on the NASDAQ stock market, but also includes companies incorporated outside the United States. These factors distinguish NASDAQ-100 from NASDAQ Composite, DJIA and S&P 500 index.

Next chapter fully describes the models from the ARCH family. The third chapter briefly presents the data used in this research. The fourth chapter presents the results of conducted research, clarifying the volatility modeling on NASDAQ-100. In final chapter, concluding remarks are given.

2. METHODS

As previously indicated, volatility is an important factor in trading. It is a crossover from the assumptions of homoscedasticity, which is necessary for application
of many time series models. The first model found for estimating volatility was ARCH (Auto Regressive Conditional Heteroscedasticity) model [3]. Since it has been introduced, there were many attempts to improve and upgrade the original ARCH model. The best known are GARCH [4], EGARCH [6] as well as GJR-GARCH [9].

ARCH is the first model that provides a systematic framework for volatility modeling [2]. An ARCH(p) model assumes that:

\[ a_t = \sigma_t \varepsilon_t \]  \hspace{1cm} (1)

\[ \sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j a_{t-j}^2 \]  \hspace{1cm} (2)

where \( \alpha_0, \alpha_p \) and \( p \) are nonnegative coefficients. Coefficient \( \varepsilon_t \) represents a sequence of independent and identically distributed (i.i.d.) random variables with zero mean and unit variance. In practice, \( \varepsilon_t \) is often assumed to follow the standardized normal distribution \( \varepsilon_t \sim N(0,1) \). By definition, \( a_t \) is a serially uncorrelated sequence, called innovation, with zero mean and the conditional variance of \( \sigma_t^2 \), which may be nonstationary. The basic idea of ARCH models is that (a) the mean corrected asset return \( a_t \) is serially uncorrelated, but dependent, and (b) the dependence of \( a_t \) can be described by a simple quadratic function of its lagged values.

Generalized Autoregressive Conditional Heteroscedasticity, GARCH(p,q), is a generalization of ARCH model by making the current conditional variance dependent on the \( p \) past squared innovations as well as the \( q \) past conditional variances [1, 2, 10]. The GARCH(p,q) model can be written as:

\[ \sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j a_{t-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \]  \hspace{1cm} (3)

where \( \alpha_0, \alpha_p, \beta_q, \ p \) and \( q \) are nonnegative coefficients. By accounting for the information in the lag(s) of the conditional variance in addition to the lagged \( a_{t-j}^2 \) terms, the GARCH model reduces the number of parameters required. Most popular and most used model is actually GARCH(1,1), which can be written as:

\[ \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \]  \hspace{1cm} (4)

The model provides a simple parametric function that can be used to describe the volatility evolution.

3. DATA CHARACTERISTICS

The data for this study consists of NASDAQ-100 stock market daily closing price index during the period from January 3, 2005 to June 30, 2011, including 1636 observations - trading days. All data were gathered from the database of Yahoo Finance website [11].

Figure 1 presents NASDAQ-100 daily close prices for the given period, and the NASDAQ-100 trading volume is given in the Figure 2. Higher volume for a security is an indicator of higher liquidity. In order to get stationary financial times series we transform prices into logarithmic returns, which are presented in Figure 3. All values match the defined period (January 2005 to June 2011).

Figure 1. NASDAQ-100 daily close prices

Figure 2. NASDAQ-100 trading volume

Figure 3. NASDAQ-100 logarithmic returns

Volatility continues to be a problem after the major shock of 2009 as the Figure 3 displays. A careful look at Figure 3 reveals that the return series have volatility clusters.

Table 1 shows the basic statistical characteristics of the return series. The kurtosis in these data suggests that
their daily return series have a fat-tailed distribution as compared with Gaussian distribution.

**Table 1.** Basic statistical characteristics for NASDAQ-100 return time series

<table>
<thead>
<tr>
<th>Statistics</th>
<th>NASDAQ-100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0002272</td>
</tr>
<tr>
<td>Median</td>
<td>0.00108</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.1111</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.1184</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.015221</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.0913</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>11.000</td>
</tr>
</tbody>
</table>

4. RESULTS

This chapter presents the results of Autocorrelations and Heteroscedasticity tests using Ljung-Box test [12] and Engle's ARCH test. Afterwards the parameter estimation is performed and discussed. All tests are conducted using MatLab software package.

4.1 Autocorrelation

As can be seen in Figure 4, autocorrelation function (ACF) for NASDAQ-100 returns becomes significant at lags 1, 2, 3, 12, 15, 16, 18 and 20 at 95% confidence level. In this case, correlation between NASDAQ-100 returns is significant.

4.2 Heteroskedasticity

In order to test heteroskedasticity, we can plot ACF for the squared returns. Figure 5 presents the image of the squared returns, which gives us the basic insight into the issue. The GARCH effects (heteroscedasticity) is also tested using Engle's ARCH test for the residuals: ARCH statistics value 525.1257 is greater than critical value 37.5662 and the null hypothesis is rejected at significance level 0.01. We accept the alternative hypothesis that there is heteroscedasticity.

4.3 Innovations

Innovations (shocks) can be interpreted as standard deviation of returns (1). The squared innovations are calculated and presented in Figure 6.

4.4 Model estimates

GARCH(1,1) parameters are estimated using MatLab software package (see Section 2, model 4), and presented in Table 2.

**Table 2.** GARCH(1,1) model estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Std. Error</th>
<th>T-Statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>2.437e-006</td>
<td>6.68e-007</td>
<td>3.6458</td>
<td>0.0003</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.074201</td>
<td>0.0093864</td>
<td>7.9052</td>
<td>0.0000</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.91248</td>
<td>0.01161</td>
<td>78.5376</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Significant at 0.01 level

As shown in Table 2, all parameters from the model (4) are shown to be significant. Also, the sum $\alpha_1 + \beta_1 = 0.986681$ is less than 1, which is required to
have a mean reverting variance process (the proposition for weak stationarity).

In order to inspect the relationship between the innovations (i.e., residuals) derived from the fitted model, the corresponding conditional standard deviations, and the observed returns, we presented the plot that compares of innovations, conditional standard deviations and returns (Figure 7).

![Figure 7. Comparison of innovations, conditional standard deviations and returns](image)

Notice that both the innovations (top plot) and the returns (bottom plot) exhibit volatility clustering.

Figure 8 plots the standardized innovations (the innovations divided by their conditional standard deviation), and it can be noticed that they appear generally stable with little clustering.

![Figure 8. Standardized innovations](image)

This can also be seen from the ACF of the squared standardized innovations plot in Figure 9. They also show no correlation. Only one critical value appears on lag 10.

To confirm this conclusion we conducted the Ljung-Box test for squared standardized innovations. The result $Q(20) = 26.4527$ is less than critical value 31.5662 and there is no reason not to accept the null hypothesis. We conclude that there is no heteroscedasticity left in the proposed model.

![Figure 9. ACF of the squared standardized innovations](image)

The correlation of the standardized innovations is also tested using Engle’s ARCH test: ARCH statistics value 28.3037 is less than critical value 37.5662 and the null hypothesis is accepted at significance level 0.01. We accept the null hypothesis that there is no heteroscedasticity.

In the pre-estimation analysis, both the Ljung-Box test and the ARCH test indicate rejection of their respective null hypotheses. This shows significant evidence in support of GARCH effects. The post-estimate analysis uses standardized innovations based on the estimated model. These same tests now indicate acceptance of their respective null hypotheses. These results confirm the explanatory power of GARCH(1,1) model.

Furthermore, we can show that our standardized residuals do not flee far from the Normal distribution (Table 3).

### Table 3. Statistical characteristics of standardized residuals

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Standardized residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0299</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.9988</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.3548</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.0454</td>
</tr>
</tbody>
</table>

4.5 Volatility forecast

In this section we forecast volatility of NASDAQ-100 for the next 10 trading days (over the period 1-15 July 2011), which is required in Basel II Accord.

The forecast standard deviation for the next day ($t+1$) is:

$$\hat{\sigma}_{t+1}^2 = 2.437e-006 + 0.074201\alpha_t^2 + 0.91248\sigma_t^2$$

where $\alpha$ and $\sigma$ are known at time $t$.

And for the following days $i = 2, \ldots, 10$ standard deviation can be forecasted as:

$$\hat{\sigma}_{t+i}^2 = \alpha_0 + (\alpha + \beta_i)\sigma_{t+i-1}^2$$

43
Table 4. Real and forecasted standard deviations

<table>
<thead>
<tr>
<th>Days</th>
<th>Real Std.</th>
<th>Forecasted Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.015221</td>
<td>0.015048</td>
</tr>
<tr>
<td>2</td>
<td>0.015217</td>
<td>0.015028</td>
</tr>
<tr>
<td>3</td>
<td>0.015212</td>
<td>0.015009</td>
</tr>
<tr>
<td>4</td>
<td>0.015211</td>
<td>0.014991</td>
</tr>
<tr>
<td>5</td>
<td>0.015207</td>
<td>0.014972</td>
</tr>
<tr>
<td>6</td>
<td>0.015209</td>
<td>0.014954</td>
</tr>
<tr>
<td>7</td>
<td>0.015206</td>
<td>0.014936</td>
</tr>
<tr>
<td>8</td>
<td>0.015202</td>
<td>0.014918</td>
</tr>
<tr>
<td>9</td>
<td>0.015200</td>
<td>0.014900</td>
</tr>
<tr>
<td>10</td>
<td>0.015199</td>
<td>0.014882</td>
</tr>
</tbody>
</table>

5. CONCLUSION

Volatility is fundamental variable in both theoretical and practical applications, owing to its central role in option pricing and risk management. Most forecasting models are unable to predict certain time series, since the main assumption for their application is the absence of volatility. NASDAQ-100 returns prove to be a financial time series characterized by heteroscedasticity. ARCH and GARCH models have been applied to a wide range of time series analyses. This paper overcomes the stated problem using the GARCH(1,1) to model volatility of the NASDAQ-100 returns time series. Volatility performance is found to be significantly improved. Thus, the input for the further financial applications of volatility of NASDAQ-100 returns is provided.

Likewise, the analysis of ARCH and GARCH models and their many extensions provides a statistical stage on which many theories of asset pricing, portfolio analysis, value at risk or index volatility can be exhibited and tested. Future research can be directed towards integration of the proposed volatility forecasting model with artificial neural networks in order to capture nonlinear patterns.

REFERENCES


