Econometric Analysis on NASDAQ 100

Financial Econometrics

Himadri Sekhar Manna [1] Sujoy Kumar Bhadra [2] Susmit Saha [3] Ayan Paul [4]

[1] MS - 1907

[2] MS - 1920

[3] MS - 1922

[4] MS - 1903

Indian Statistical Institute, Kolkata

June 12, 2021

Outline

Outline

Outline

Outline

Data Description

- The NASDAQ 100 is a stock market index made up of 102 equity securities issued by 100 of the largest non-financial companies listed on the NASDAQ stock market. The stocks' weights in the index are based on their market capitalization, with certain rules capping the influence of the largest components. It is based on exchange, and it does not have any financial companies.
- For our analysis, we will work with the data of NASDAQ 100 starting from 2^{nd} January 2010 to 31^{st} December 2019.

Data Visualisation

Test of Stationarity of NASDAQ Data

- We use Augmented Dickey-Fuller Test for testing stationarity.
- Here, the null hypothesis is-

$$H_0 : \textit{The data is not stationary}$$

  and the alternative hypothesis is-

$$H_1 : \textit{The data is stationary}$$

- The value of the test statistic is -1.7237 and the p-value of the test is 0.6953.

Transformation for Dealing with Non-stationarity

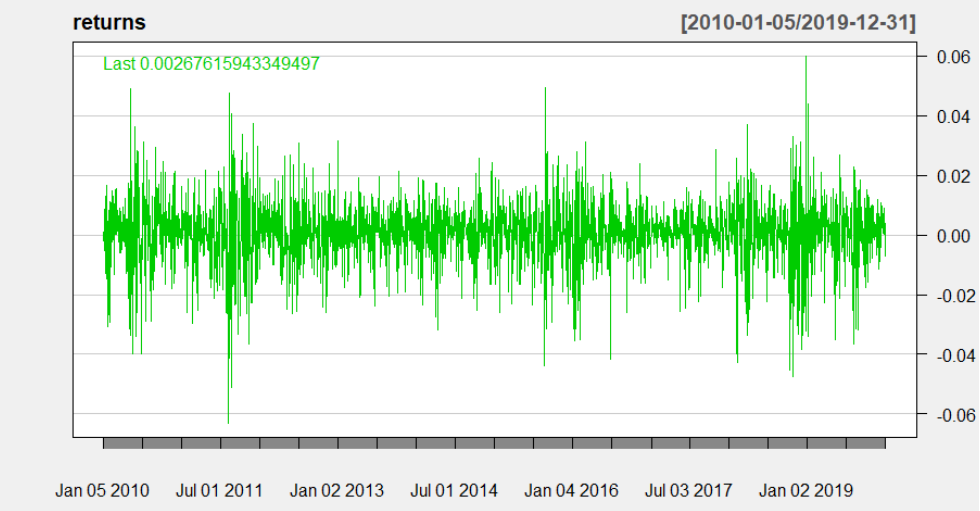
- We will work with the log-returns in our analysis, i.e.,

$$r_{t+1} = \log(\frac{S_{t+1}}{S_t})$$

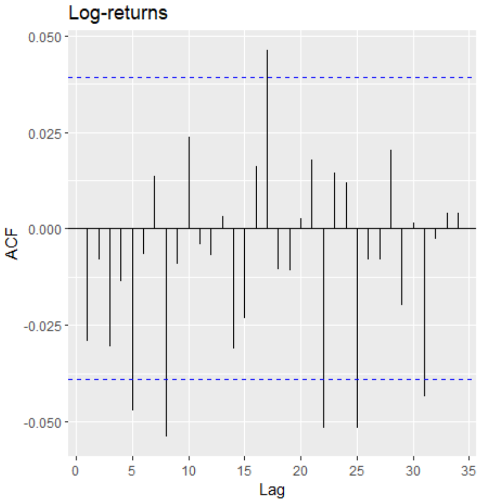
- The p-value corresponding to log-returns in Augmented Dickey-Fuller test is less than 0.01.

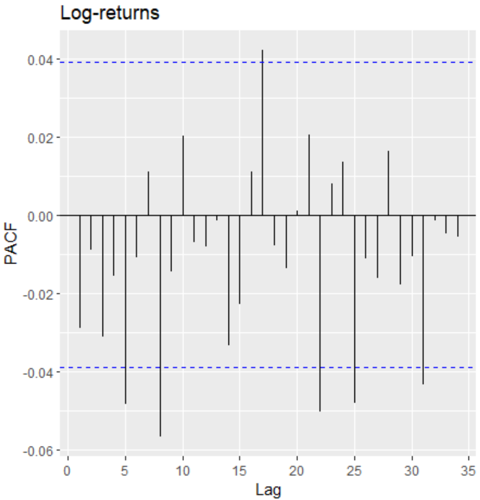
- Under 5% level of significance, we can conclude that the series of log-returns is stationary.

Plot of Log-Returns of NASDAQ



ACF and PACF Corresponding to Log-returns





ARMA Model

- In the time series, it is desirable to allow the dependent variable to be influenced by the past values of the independent variables and possibly by its own past values. Also, this kind of formulation leads to an efficient forecasting procedure.

- ARMA models is one of the most useful modelling technique in this regard. Mathematically speaking, a time series $\{x_t : t = 0, \pm 1, \pm 2, \dots\}$ is called $\mathbf{ARMA}(p, q)$ if it is stationary and

$$x_t = \sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{i=1}^{q} \theta_i w_{t-i}$$

where, $\{w_t : t = 0, \pm 1, \pm 2, \dots\}$ is *Gaussian white noise* sequence. Also, we have, $\phi_p \neq 0$, $\theta_q \neq 0$ and $\sigma_w^2 > 0$.

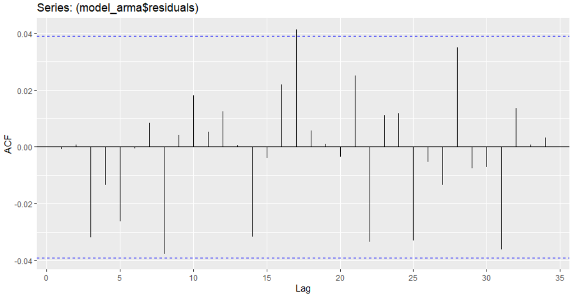
- Here, the parameters p and q are called the *auto-regressive order* and *moving average order* respectively.

ARMA Model

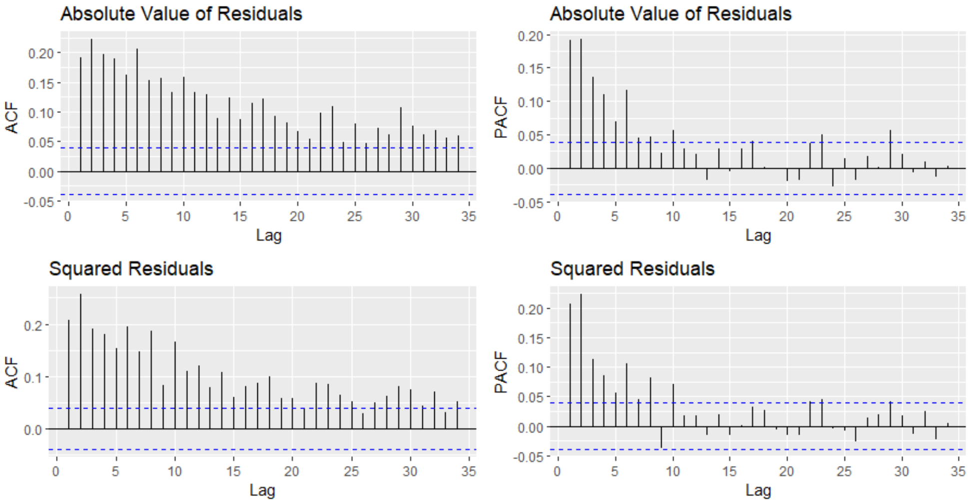
- We have used AIC as our model selection criterion.
- Keeping the idea of model-parsimony in mind, the selected ARMA model is,

$$r_t = 0.3599r_{t-1} - 0.3845r_{t-2} + 0.9805r_{t-3} - 0.3942e_{t-1} + 0.3854e_{t-2} - 1.0024e_{t-3} + 0.0223e_{t-4} - 0.0111e_{t-5} + 0.0006$$

- The p-value corresponding to **Ljung-Box** test(lag $= 30$) for the residuals is 0.473.
- Under 5% level of significance the residuals are independent.



Graphical Evidence for Presence of Auto Regressive Conditional Heteroscedasticity



LM test: Test for the Presence of ARCH

- The ACF and PACF plots show evidence of presence of Auto Regressive Conditional Heteroscedasticity. So, we have to perform a testing of hypothesis in this regard. *LM test statistics* derived by Engle is one of the useful tool to serve this purpose.

- In **ARCH** modelling, we assume that the residual $\varepsilon_t \mid \{y_{t-1}, x_{t-1}, \ldots\} \sim \mathcal{N}(0, h_t)$ where, we model h_t as, $h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$; with $\alpha_0 > 0$ and $\alpha_i \geqslant 0$ for all $i = 1, \ldots, q$. So, our null hypothesis would be, $\mathbf{H}_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_q = 0$.

- The LM statistics is product of number of observations (T) and coefficient of multiple determination from the regression of $\hat{\varepsilon}_t^2$ on a constant and $\hat{\varepsilon}_{t-1}^2, \ldots, \hat{\varepsilon}_{t-q}^2$ (R); $\hat{\varepsilon}_t$'s being the **OLS** residuals of the original model, i.e., $y_t = \boldsymbol{x}_t^T \boldsymbol{\beta} + \varepsilon$.

- It has been derived that under $\mathbf{H}_0$, this statistics follows χ_q^2 distribution. Also, we reject $\mathbf{H}_0$ for too large value of the LM statistics.

- In our case, p-values corresponding to the LM test statistic are all less than 0.05. Hence, Auto Regressive Conditional Heteroscedasticity is present in the data.

GARCH modelling

- In **ARCH** modelling, we model the conditional variance h_t as only the function of residuals of the past. A natural extension of modelling would be to incorporate previous realization of the conditional variances. **GARCH** modelling exactly do the same thing. Also, it gives a parsimonious representation of a high order ARCH process.

- A time series $\{x_t : t = 0, \pm 1, \pm 2, \ldots\}$ is said to follow **GARCH**(p, q) if the conditional variance h_t can be modelled as,

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

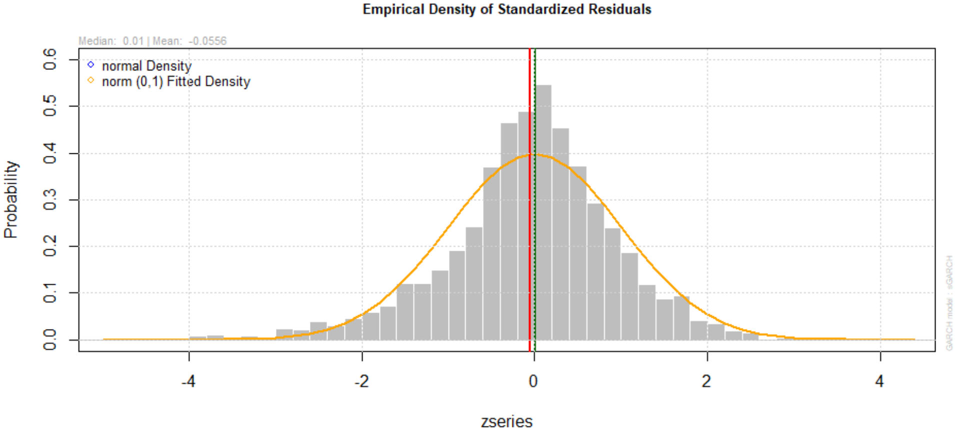
where, we impose $\alpha_0 > 0$, $\alpha_i \geqslant 0 \ \forall i = 1, \ldots, q$, $\beta_j \geqslant 0 \ \forall j = 1, \ldots, p$ to ensure strict positiveness of the conditional variance.

GARCH Model Selection

- We have used AIC as our model selection criterion.
- The selected order of GARCH is (1,1).
- The model is-

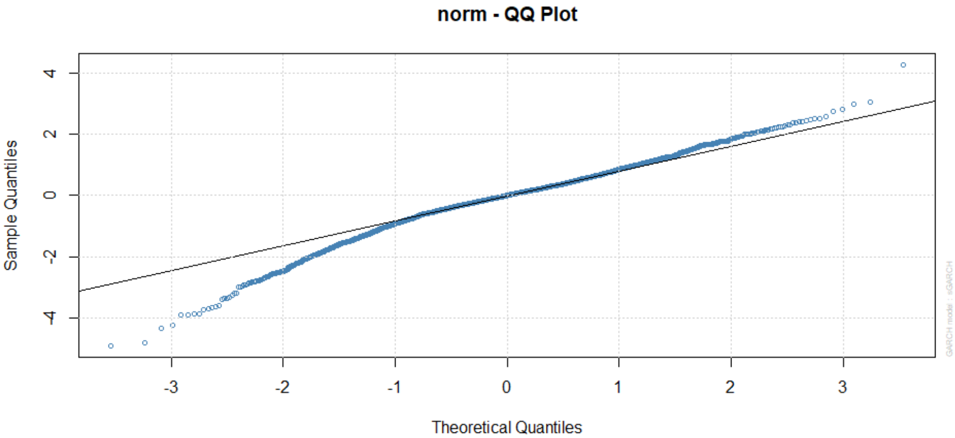
$$h_t = 0.000005 + 0.127327\varepsilon_{t-1}^2 + 0.826246h_{t-1}$$

Residual Analysis of ARMA(3,5)+GARCH(1,1)



- The histogram of the residuals deviates a little from the standard normal density.

Residual Analysis of ARMA(3,5)+GARCH(1,1)

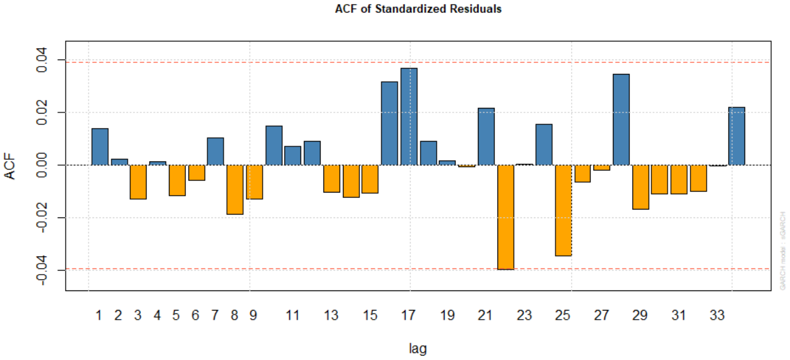


- The qqplot shows evidence towards non-normality.

Residual Analysis of ARMA(3,5)+GARCH(1,1)

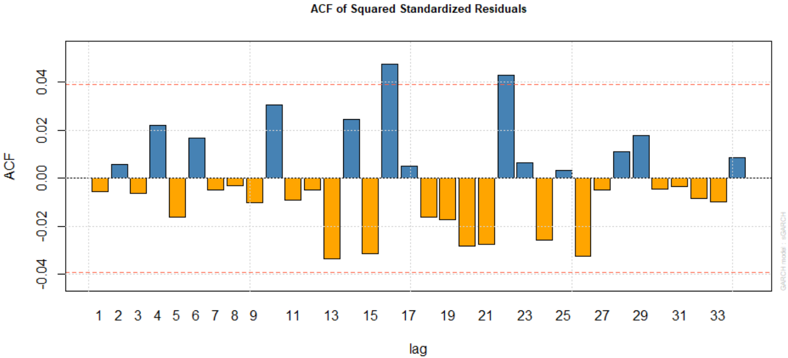
- To deal with the non-normality, we considered transformed prices. Here, we considered **Box-Cox** Transformation.
- Even considering **Box-Cox** Transformation we observed quite similar QQ-plot. Hence, we worked with the original prices.

Residual Analysis of ARMA(3,5)+GARCH(1,1)



- There is no significant dependency amongst the standardized residuals.
- The p-value corresponding to Ljung-Box Test for the standardized residuals also shows evidence towards no serial correlation.
- Hence, all significant part is explained by our model.

Residual Analysis of ARMA(3,5)+GARCH(1,1)



- There is no significant dependency amongst the squared standardized residuals.
- The p-value corresponding to Ljung-Box Test for the squared standardized residuals also shows evidence towards no serial correlation.
- Hence, there is no unexplained ARCH effect.

Forecasting

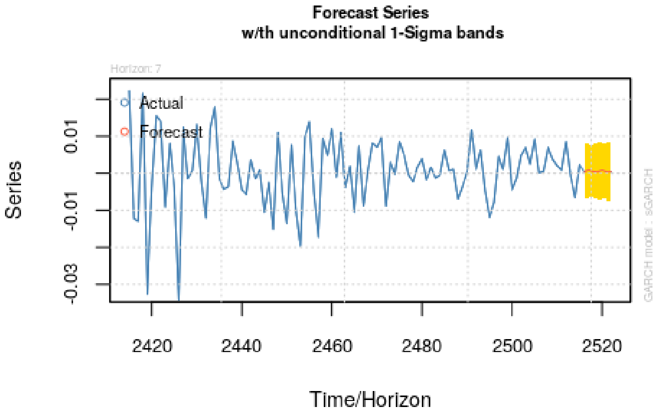


Figure: Forecasting using GARCH modelling

EGARCH

- One of the main limitations of GARCH is that the conditional variance h_t is symmetric in the lagged ε_t's. But, it is often found that this consequence of GARCH is inappropriate for modelling the volatility of the returns on stocks.
- To tackle this limitation, Nelson proposed exponential GARCH(p, q) where h_t can be modelled as,

$$\log(h_t) = \alpha_0 + \sum_{i=1}^{q} \alpha_i g(\eta_{t-i}) + \sum_{j=1}^{p} \beta_j \log(h_{t-j})$$

where, $g(\eta_t) = \theta\eta_t + \gamma(|\eta_t| - \mathbb{E}(|\eta_t|))$; also here η_t serves as the forcing variable for both the conditional variance and the error.

EGARCH Model Selection

- We have used AIC as our model selection criterion.
- Keeping model-parsimony in mind we selected order of EGARCH as (4,4).
- The model is-

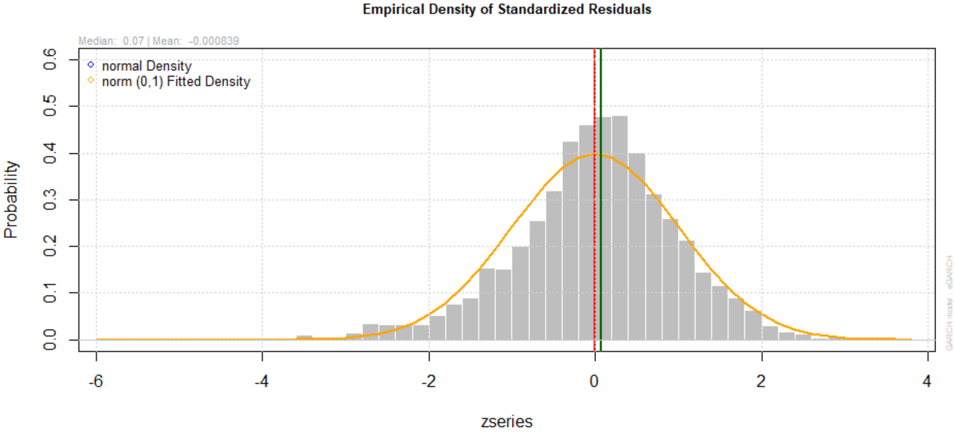
$$\log(h_t) = \alpha_0 + \sum_{i=1}^{4} \alpha_i g(\eta_{t-i}) + \sum_{j=1}^{4} \beta_j \log(h_{t-j})$$

$$\omega = -0.071621,\ \alpha_1 = -0.230828,\ \alpha_2 = -0.068536,\ \alpha_3 = 0.091545,\ \alpha_4 = 0.207841$$

$$\beta_1 = 0.563637,\ \beta_2 = 0.625449,\ \beta_3 = 0.696933,\ \beta_4 = -0.893837$$

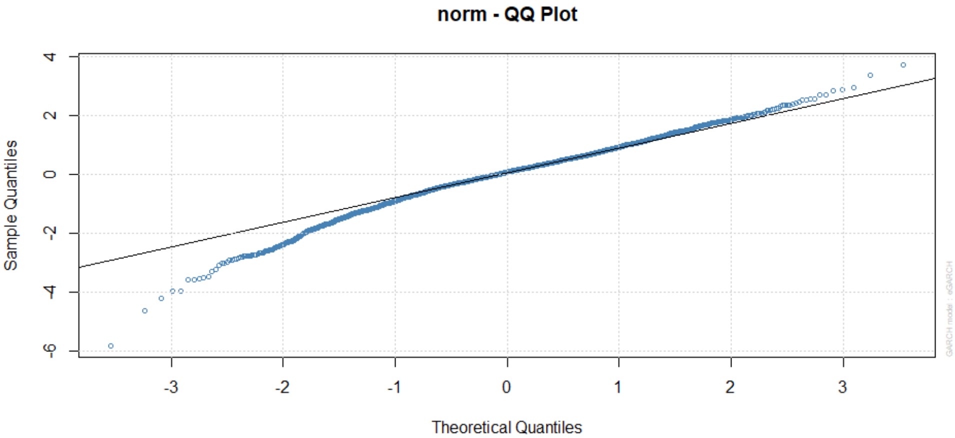
$$\gamma = 0.091800$$

Residual Analysis of ARMA(3,5)+EGARCH(4,4)



- The histogram of the residuals deviates a little from the standard normal density.

Residual Analysis of ARMA(3,5)+EGARCH(4,4)

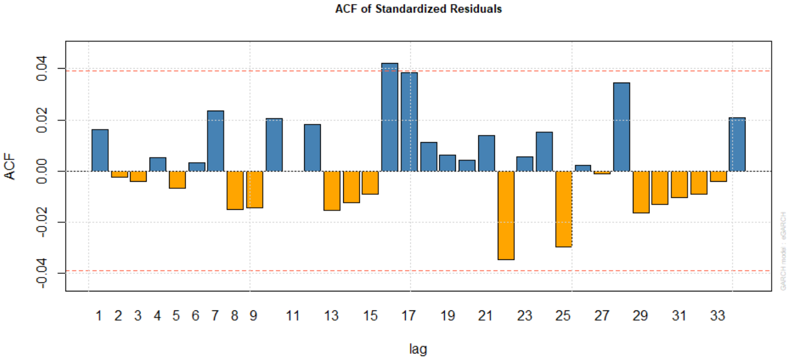


- The qqplot shows evidence towards non-normality.

Residual Analysis of ARMA(3,5)+EGARCH(4,4)

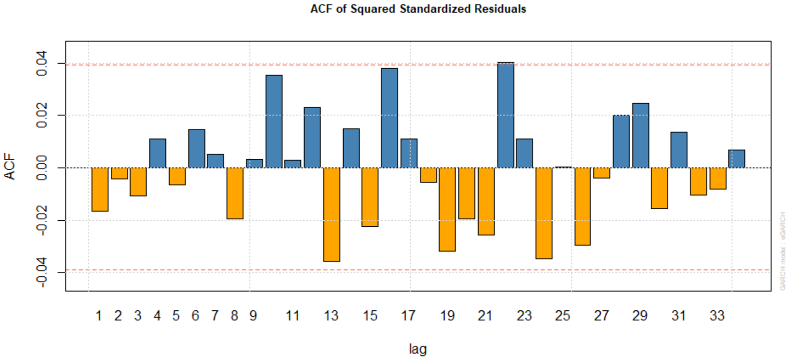
- To deal with the non-normality, we considered transformed prices. Here, we considered **Box-Cox** Transformation.
- Even considering **Box-Cox** Transformation we observed quite similar QQ-plot. Hence, we worked with the original prices.

Residual Analysis of ARMA(3,5)+EGARCH(4,4)



- There is no significant dependency amongst the standardized residuals.
- The p-value corresponding to Ljung-Box Test for the standardized residuals also shows evidence towards no serial correlation.
- Hence, all significant part is explained by our model.

Residual Analysis of ARMA(3,5)+EGARCH(4,4)



- There is no significant dependency amongst the squared standardized residuals.
- The p-value corresponding to Ljung-Box Test for the squared standardized residuals also shows evidence towards no serial correlation.
- Hence, there is no unexplained ARCH effect.

Forecasting

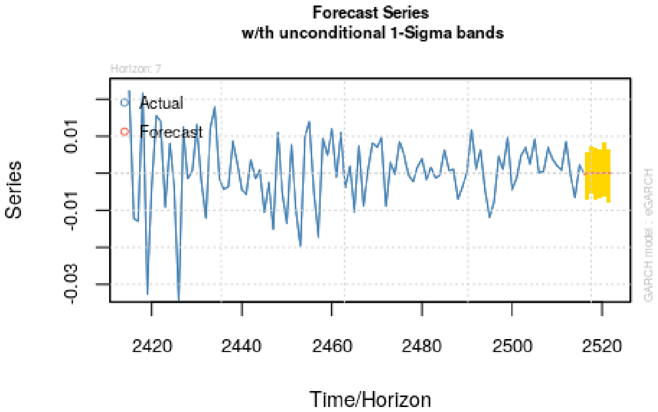


Figure: Forecasting using EGARCH

Prediction

To compare the prediction performance of the models, we will take a look at the closing price predicted by the model for $T = t+1,\ t+2,\ t+3$ and $t+4$, where the model is built based on t time-points.

	NASDAQ Closing Price	ARMA Forecast	ARMA+GARCH Forecast	ARMA+EGARCH Forecast
02-01-2020	8872.22	8686.217	8687.271	8679.56
03-01-2020	8793.9	8681.749	8692.89	8678.853
06-01-2020	8848.52	8680.024	8696.217	8680.203
07-01-2020	8846.45	8687.204	8702.345	8680.864

Comments

- The ARMA+GARCH model is the best model for forecasting the NASDAQ 100 price. Considering the ARCH effect has resulted in better forecasting.
- The ARMA+EGARCH model is not performing better then ARMA+GARCH model in terms of forecasting. The possible reason might be the absence of **Leverage Effect** for NASDAQ 100 within the time-period of interest.