**NATIONAL COLLEGE OF IRELAND**

**SCHOOL OF COMPUTING**

**MASTERS IN FINTECH**

**FORECASTING AND MODELLING STOCK MARKET VOLATILITY: EVIDENCE FROM A DEVELOPED**

**STOCK MARKET INDEX**

**SUBMITTED**

**BY**

**SEYINGBO OLUWAGBENGA ABAYOMI**

**ABSTRACT: *The importance of modelling and forecasting the volatility of stock market returns cannot be overemphasized because efficient information about volatility of the stock market returns enables investors to make accurate financial decisions which involves risk management, option pricing, asset allocation, asset pricing, portfolio management and hedging strategy. The models for forecasting stock market volatility have been evolving over time since the introduction of Autoregressive Conditional Heteroscedasticity (ARCH), several studies have also introduced other models such GARCH, EGACRH, TGARCH, PGARCH, GJR GARCH. However, the findings of these models have been inconsistent. This is because researches done on stock market are carried out on different stock market index, time frame, and most importantly, the stock markets are faced with different exposure such as macroeconomic instability, political instability, external shocks, capital flight etc which also vary from one country to another. It is against this setback that this study modelled and forecast stock market volatility using developed stock market index such as standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100. The study adopted symmetry and asymmetry GARCH models such GARCH, EGARCH, TGRACH and PGARCH in forecasting stock market volatility, while Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percent Error (MAPE) were used to compare the forecasting ability of these models.***

1. **INTRODUCTION**

Forecasting and modelling stock market volatility has pulled in a great deal of worry for different stakeholders, investors, regulators and institutions. This concern was triggered by the stock market crash that

occur in 1987. The importance of forecasting the volatility of stock returns cannot be overemphasized because investors uses the estimate of expected returns, risk and volatilities as a yardstick before making crucial financial decisions such as risk and portfolio management, asset pricing, asset allocation and hedging strategy. (Bollerslev, Pattona and Quaedvlieg, 2016).

The models for forecasting stock market volatility have been evolving over time since the introduction of Autoregressive Conditional Heteroscedasticity (ARCH) by Engle in 1982 and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model by Bollerslev in 1986. These models usually use historical

information or past occurrence such as previous prices or previous stock returns for predicting financial times series which possess the attributes of heteroscedasticity. However, ARCH and GARCH models does not estimates volatility asymmetric. Volatility asymmetry occurs when stock return is higher when the stock market is performing badly than when the market is performing well.

This impediment prompted further improvement of other GARCH models such Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) model which was presented by Nelson in 1991 and Non-Linear Asymmetric Generalized Autoregressive Conditional Heteroscedasticity (NGARCH) model which was adopted by Higgins and Bera in 1992 (Hajizadeh, Seifi, Zarandi and Turksen, 2012). Subsequently, several models such as Threshold GARCH (TGARCH), Integrated Generalized Autoregressive Conditional heteroskedasticity (IGARCH), GARCH-in-mean (GARCH-M), Quadratic GARCH (QGARCH) and Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) were also introduced and each of this model has its own limitations and advantages. (Sentana, 1995; Glosten, Jagannathan and Runkle, 1993; Zakoian, 1994).

The investigation of Wei (2012) detailed that Quadratic GARCH (QGARCH) estimates the best predicting ability of stock market volatility than GARCH model, while, Yeh and Lee (2000) uncovered that GJR-GARCH model performs better than GARCH models, thus, it estimates the best forecasting ability of the Chinese stock market. Furthermore, Awartani and Corradi (2005) discovered that asymmetry GARCH models is superior than symmetry GARCH model in terms of forecasting ability, while, the discoveries of McMillan, Speight and Apgwilym (2000) additionally found that GARCH model gives the better forecast of stock market volatility than the EGARCH model and this discovery repudiates the findings of Awartani and Corradi (2005).

The inconsistency is because researches done on stock market are carried out on different stock market index, time frame, and most importantly, stock markets are faced with different exposure such as macroeconomic instability, political instability, external shocks, capital flight etc which also vary from one country to another.

Generally, stock markets are usually influenced by good or bad news, thereby, leading to positive or negative shocks, international capital flight, leverage effect, volatility clustering, conditional volatility, time varying volatility and leptokurtosis. As indicated by Black (1976), leverage effect emerges when changes in stock returns are negatively associated with changes in volatility while, volatility clustering emerges when huge changes of stock returns are trailed by large volatilities and vice versa. It is based on this issues that this study sets to achieve the following research objectives:

(i). to inspect the presence of volatility clustering on the standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100; (ii). to investigates the leverage effect of volatility on standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100; (iii). to investigates the presence and the effect of conditional volatility and time varying volatility on standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100; (iv). to estimates and compares the volatility clustering, conditional volatility, time varying volatility, leptokurtosis, leverage effect of standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100; and (v). to compare the forecasting ability of symmetry and asymmetry GARCH models on standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100.

It is on this premise that this study extract information about the variance of future returns from historical data on Standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100 using GARCH-type models. In addition, the forecasting ability and the performance of these models were evaluated and compared through the error measurement estimates such as Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percent Error (MAPE).

The remaining study is organized as follows: section two explains the state of art on models on stock market volatility. Section three critically review studies related stock market volatility using developed and emerging stock market index as a point of reference. Section three lay emphasis on methodology, model specification and model evaluation. Section four explains the Apriori expectations of the study on based on theory and other related studies while, section five gives a conclusion and future work.

1. **STATE OF ART**

The investigation of Qifa, Zhongpu, Cuixia, and Yezheng (2019) clarified that there are three models of estimating volatility. The first model is known as the GARCH models and Stochastic Volatility (SV). These models use past information or trend of stock returns for predicting future stock returns and volatility. GARCH models are used to test the stock market efficiency, value at risk. These models also explain excess kurtosis and volatility clustering of stock market indexes. The second model is known as Implied Volatility (IV), and this model is determined from an option pricing model, while the third model is Realized Volatility (RV), which uses high frequency data for modelling stock market volatility (Lin, 2018).

Taylor (1986) introduced a Stochastic Volatility (SV) model to predict stock market volatility, and from that point forward, different studies have widely utilized the model to assesses and forecast stock market volatility. However, this model does not capture the entire day data or information on stock returns, for example, intraday stock market returns and volatilities, because it uses low-frequency data (Gong and Lin, 2019). It is against this setback that Andersen and Bollerslev (1998) introduced another model known as Realized Volatility (RV) that uses high-frequency data for measuring volatility.

The investigation of Gong and Lin (2019) clarified that since the introduction of Realized Volatility (RV) by Andersen and Bollerslev in 1998, the model has attracted a lot of interest in the body of literature because its prediction were accurate than other models used in measuring stock market volatility In a similar vein, Hajizadeh, Seifi, Zarandi and Turksen (2012) investigated a hybrid modelling approach for forecasting the volatility of Standard & Poor's 500 index using two hybrid models such as EGARCH and Artificial Neural Networks. The findings revealed that the hybrid model provides a better forecast of stock market volatility. The study of Hamid and Iqbal (2004) used Artificial Neural Network (ANN) to forecast the volatility of Standard & Poor’s 500 index, and it was revealed that ANNs are superior to implied volatility in forecasting stock market volatility. However, Chen, Zhou, and Dai (2015) revealed that long short-term memory (LSTM) model is superior than feedforward neural network model in forecasting Chinese stock market volatility.

Furthermore, Tseng, Cheng, Wang, and Peng (2008) introduced a hybrid model by combining EGARCH model and feedforward neural network to evaluate the volatility of Taiwan stock market index. The findings revealed that the hybrid model is superior to EGARCH model in forecasting the volatility of Taiwan stock market index. In addition, Hajizadeh, Seifi, Zarandi, and Turksen (2012) also developed a hybrid model by integrating EGARCH and feedforward network model. The study discovered that the hybrid model is superior than the ANN model in forecasting stock market volatility. Subsequently, the study of Hernández (2017) evaluates and compared the forecasting ability of GARCH-type models and hybrid neural network models and it was discovered that the hybrid neural networks performs better than the GARCH models.

1. **RELATED WORK**

The investigation on stock market volatility is wide, and accordingly, it has created a great deal of issues in the scholarly world. Several studies have examined stock market volatility in developed and emerging economies. The investigation of Ching and Siok (2013) compares the performance of GARCH-type models on stock market volatility using Malaysian stock market index as a case study. The study employed GARCH-type models such as symmetry and asymmetry GARCH models and three statistical error measures tools such as mean squared error, root means squared error and mean absolute percentage were used to measure and compare the forecasting ability of the stock market volatility. The study covers pre-crisis, during crisis and post crisis period. The findings revealed that symmetric and asymmetric GARCH models have different performances in different time frames. The study further revealed that symmetric GARCH model perform better than the asymmetric GARCH before the crisis period while asymmetry GARCH perform better than the symmetry GARCH during the crisis period.

The study of Lin (2008) used GARCH models to Forecasts the China’s Stock Market Volatility using SSE Composite index as a case study. The study reveals that the stock market shows the presence of time-varying, volatility clustering, leptokurtosis distribution, ARCH and GARCH. In addition, the EGARCH (1, 1) model is superior to other models such as GARCH (1, 1) and TARCH (1, 1) in forecasting stock market volatility.

The study of Wong and Kok (2005) examines the forecasting ability of ARCH and GARCH models on ASEAN stock markets using Malaysia, Singapore, Thailand, Indonesia and Philippines stock market as a case study. The findings revealed that the ARCH-M model is superior than GARCH model in forecasting the volatility of Malaysia, Singapore and Thailand stock market, while random walk model is superior than ARCH and GARCH models in forecasting the volatility of Indonesia and Philippines stock market. The study further revealed that, the TGARCH and EGARCH models are superior in forecasting stock market volatility during the post-crisis period.

The study of Omar and Halim (2015) examined the volatility of Malaysian Stock market with main objective to investigate the behaviour of stock return volatility. The study employed the use of three GARCH model such as GARCH (1,1), EGARCH and TGARCH as its estimation technique and it was revealed that there is presence of volatility clustering, leverage and persistence effects on the stock market volatility. The study also revealed that EGARCH model is superior than other GARCH model in forecasting the Malaysian stock market.

1. **METHODOLOGY**

This study proposes to adopt the KDD approach. The diagram below depicts the step by step approach towards data mining, interpretation and evaluation.

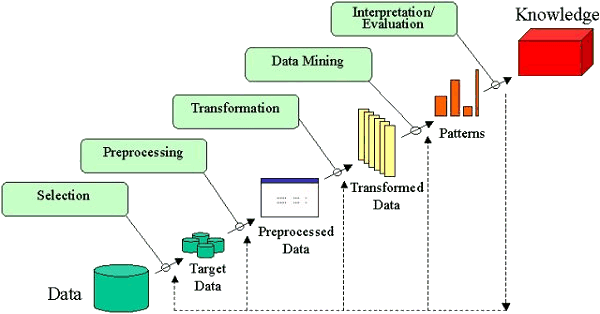


Figure 1 KDD

The study adopted the GARCH type models which consist of GARCH (p, q), TGARCH (p, q), EGARCH (p, q) and PGARCH (p, q).

(i). ARCH Model: ARCH (q) model estimates the future variance of a times series using the information from the previous variance. The “q” signifies stands order of the previous variance. The assumption of the ARCH model is that the current and future prices is usually determined by its previous prices. Thus, it follows a random walk theory. The ARCH Model consists of the mean equation and the variance equation.

(a). Mean Equation

Y = λ0 + λ1X𝑡 + µ𝑡……………………….…………….. (1)

(b). Variance Equation

= ∝0 + Σ α1 μ …………………….. (2)

The variance equation possesses the attributes which enables the model to forecast volatility.

Where:

∂ = current volatility

α1 = parameter measuring the effect of its lagged value of μ

(ii). GARCH Model: The GARCH (p, q) improves on the limitations of the ARCH (q) model because it consists of an order of previous conditional variance and previous residual in determining conditional variance. The GARCH model is specified in order (*p, q*).

𝜎 = 𝜆o + Σ 𝛼𝑖 ε + Σ 𝛽𝑗 𝜎 ………... (3)

Where:

𝜎𝑡2 = current volatility,

𝛼𝑖 = parameter measuring the effect of previous residual of 𝜀

𝛽𝑗 measures the effect of change in its lagged value, of 𝜎.

(iii). TGARCH Model: TGARCH (p, q), EGARCH (p, q) and PGARCH (p, q) improves on the limitations of GARCH (p, q) model which failed to estimate the asymmetry effect of stock market volatility. The “p” is the order of the previous residual term while the “q” remains the order of the previous conditional variance. The TGARCH looks in the order of *p, q, d.* The *p, q* and *d* stand for the number of ARCH, GARCH and asymmetric terms.

𝜎 = 𝜆o + Σ𝛼𝑖 μ+ Σ𝛽𝑗 γiIt-1 μ 𝜎2⁄(t-j) +Σ𝛽𝑗 𝜎 …………. (4)

Where:

𝐼t-1=1 if 𝜇 <0 and 0 otherwise.

(iv) EGARCH Model: The EGARCH model measures the effect of good news or bad news on volatility.

Log (𝜎) = 𝜆0 Σ{𝛼𝑖 **||**+ +𝛾𝑖 ()} + Σ𝛽𝑗log (𝜎) ………………………………(5)

(v) PGARCH model: The PGACRCH also measure the asymmetric effects of a volatility and it is stated as follows:

𝜎 = 𝜆0 + Σ𝛼𝑖 (|𝜀𝑡−1|+𝛾𝑖𝜀𝑡−1)d + Σ 𝛽𝑗𝜎………………………………………….. (6)

**Estimation Techniques: The conditional maximum likelihood method**

According to Asteriou and Hall (2007), the ARCH models was developed based on the assumptions of maximum-likelihood approach. The conditional maximum likelihood method Is also the basis upon which all other GARCH models such as ARCH, GARCH, TGARCH, EGARCH and PGARCH. The following are the steps in building ARCH and GARCH models:

(i). check the stationarity of the time series or variable to be estimated.

(ii). test for the presence of heteroscedasticity, ARCH effect and clustering volatility.

(iii). conduct a diagnostic check.

**Data and Sources of Data**

The data proposed to be used for this study consists of Standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100. This study covers the period between 2015 to 2018. The justification for the choice of this period is because of stock market sell-off that occurs in 2015 and the Brexit vote in 2016. The stock market index experienced a global decline in the value of stock prices because of stock market sell-off and the good and bad news associated with Brexit vote.

The data used for this study will be sourced from Quandl and Yahoo Finance which consist of 4600 observations. The sample of the data will be divided into two such as in-sample and out-sample. The in-sample data will be used to estimates and compares the volatility clustering, time varying volatility, leptokurtosis, leverage effect, conditional volatility on the standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100, while the out-sample data will used to compare the forecasting ability of symmetry and asymmetry GARCH models on standard and Poor’s 500, Dow Jones Industrial Average, Financial Times Stock Exchange (FTSE) 100.

Furthermore, the data will be analyzed through the KDD approach which consist of data selection, preprocessing, data transformation, data mining, interpretation, evaluation and generation of knowledge.

**Model Evaluations**

The study proposes to apply three error measures to estimates the forecasting ability of ARCH, GARCH, TGARCH, EGARCH and PGARCH. These error measurement estimators are Mean Squared Error (MSE), Root Means Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

**(i). Mean Squared Error (MSE)**: Mean Squared Error (MSE) is usually referred to as the best error measurement to determine which volatility models avoid huge errors with precision and accuracy. The Mean Squared Error (MSE) is stated as follows:

MSE =

Where:

ei = yt - ˆyt

yt = actual observed value at time t

ˆyt = fitted value at time t.

(ii). **Root Means Squared Error (RMSE):** The Root Means Squared Error RMSE is also one of the error measurement estimators widely used by several studies. The RMSE is stated as follow:

RMSE =

Where:

ei = yt - ˆyt

yt = actual observation value at time t

ˆyt = fitted value at time t.

(iii). **Mean Absolute Percentage Error (MAPE):** Mean Absolute Percentage Error (MAPE) is stated as follows:

MAPE = \* 100

n

1. **APRIORI-EXPECTATION**

The study of Lin (2018) used GARCH model to forecast stock market volatility using SSE Composite Index as a case study and it was revealed that time-varying and clustering are one of the characteristics of SSE Composite Index and EGARCH (1,1) shows a better forecasting ability than other GARCH models. Hajizadeh, Seifi, Zarandi and Turksen (2012) investigated a hybrid modelling approach by combining EGARCH model and ANN to forecast the volatility of Standard & Poor's 500 and it was discovered that the hybrid model provides a better forecast of stock market volatility.

The study of Hamid and Iqbal (2004) used Artificial Neural Network (ANN) to forecast the volatility of Standard & Poor’s 500 index, and it was revealed that ANNs are superior to implied volatility in forecasting stock market volatility. In the same vain, Chen, Zhou, and Dai (2015) revealed that long short-term memory (LSTM) model is superior than feedforward neural network model in forecasting Chinese stock market volatility.

In addition, Tseng, Cheng, Wang, and Peng (2008) built a hybrid model by combining EGARCH model and feedforward neural network to evaluate the volatility of Taiwan stock market index and it was also revealed that the hybrid model is superior to EGARCH model in forecasting the volatility of Taiwan stock market index. Hernández (2017) evaluates and compared the forecasting ability of GARCH-type models and hybrid neural network models and it was discovered that the hybrid neural networks performs better than the GARCH models.

1. **CONCLUSION AND FUTURE WORK**

The importance of modelling and forecasting the volatility of stock market returns cannot be overemphasized because the estimates of future returns and volatility assist the investors to make accurate financial decisions which involves risk management, option pricing, asset allocation, asset pricing, portfolio management and hedging strategy. The models for forecasting stock market volatility have been evolving over time since the introduction of Autoregressive Conditional Heteroscedasticity (ARCH), several studies have also introduced other models such GARCH, EGACRH, TGARCH, PGARCH, GJR GARCH.

However, the findings of these models have been inconsistent because several studies are carried out on different stock market index and time frame. The study adopted symmetry and asymmetry GARCH models such GARCH, EGARCH, TGRACH and PGARCH in forecasting stock market volatility, while Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percent Error (MAPE) were used to compare the forecasting ability of these models.

Further studies should capture and compare other stock market index of developed and emerging economies such as NASDAQ Composite, Euronext, Nikkei 225, Russell 2000 index, NASDAQ 100 using the GARCH models and time frame to estimate stock market volatility.

1. **REFERENCES**

Alberga, D., Shalita, H., and Yosef, R. (2008). Estimating stock market volatility using asymmetric GARCH models, *Applied Financial Economics*, 18:1201-1208.

Andersen, T. G., and Bollerslev, T. (1998). Answering the critics: Yes, ARCH models do provide good volatility forecasts, *Int. Econ. Rev.* 4: 885–905.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2003). Modeling and forecasting realized volatility, *Econometrica*, 71, 579–625.

Asteriou, D. and Hall, S. G. (2007). *Applied Econometrics: A Modern Approach.* New York. Palgrave Macmillian. Revised Edition.

Awartani, B. M. and Corradi, V. (2005). Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries, *International Journal of Forecasting*, 21: 167-183.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, 31, 307–327.

Bollerslev, T., Pattona, A., and Quaedvlieg, R. (2012). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192, 1–18.

Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. In Proceedings of the 2015 IEEE international conference on big data, 2823–2824.

Ching, M. L. and Siok, K. S. (2013). Comparing the performances of GARCH-type models in capturing the stock market volatility in Malaysia, *International Conference on Applied Economics (ICOAE). Procedia Economics and Finance* 5: 478 – 487.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance UK inflation. *Econometrica,* 50, 987–1008.

Engle, R. and Ng, K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance*, 48 (5): 1022–1082

Glosten, L., Jagannathan, R., and Runkle, D. (1993). On the relation between the expected value and the volatility nominal excess return on stocks. *Journal of Finance*, 46, 1779–1801.

Gong, X. and Lin, B. (2019). Modelling Stock Market Volatility using new HAR-type models. *Physica A*, 516: 194–211.

Hajizadeh E., Seifi A., Zarandi, M. H., and Turksen, I.B. (2012). A hybrid modeling approach for forecasting the volatility of S&P 500 index return. *Expert Systems with Applications*, 39:431–436

Hamid, S. A., & Iqbal, Z. (2004). Using Neural Networks for Forecasting Volatility of S&P 500 Index futures prices. *Journal of Business Research*, 57(10):1116–1125.

Hernández, E. (2017). Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors. *Expert Systems with Applications*, 84, 290–300.

Higgins, M. L. and Bera, A. K. (1992) A Class of Non-Linear ARCH models *Int. Econ. Rev*. 33, 137-158.

Lin, Z. (2018). Modeling and Forecasting the Stock Market Volatility of SSE Composite Index using GARCH models. *Future Generation Computer Systems*, 79 (2018): 960–972.

McMillan, D. G., Speight, A. E. H., & Apgwilym, O. (2000). Forecasting UK Stock Market Volatility. *Applied Financial Economics*, 10, 435–448.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: *A New Approach*. Econometrica, 59, 347–370.

Omar, N. A and Halim, F. A. (2015). Modelling Volatility of Malaysian Stock Market using GARCH Models. *International Symposium on Mathematical Sciences and Computing Research (iSMSC), 447- 452.*

Qifa, X., Zhongpu B., Cuixia, J., and Yezheng, L. (2019). Does Google search index really help predicting stock market volatility? Evidence from a modified mixed data sampling model on volatility. *Knowledge-Based Systems* 166, 170–185

Sentana, E. (1995). Quadratic ARCH Models, *The Review of Economic Studies*, 62, 639-661.

Tseng, C. H., Cheng, S. T., Wang, Y. H., & Peng, J. T. (2008). Artificial Neural Network Model of the Hybrid EGARCH Volatility of the Taiwan stock index option prices. *Physica A*: Statistical Mechanics and its Applications, 387(13), 3192–3200.

Wei, W. (2002). Forecasting stock market volatility with non-linear GARCH models: A case for China. *Applied Economics Letters*, 9, 163–166.

Wong, Y. C., Kok, K. L. 2005. A Comparison of Forecasting Models for ASEAN Equity Markets. *Sunway Academic Journal*. 2, 1.

Yeh, Y. and Lee, T. (2000). Interaction and volatility asymmetry of unexpected returns in the greater china stock markets, *Glob. Finance J.* (11): 129–149.

Zakoian, J. (1994). Threshold heteroskedastic models, *Journal of Economic Dynamics and Control*, 18:931-955.