2021

Modelling and forecasting of crude oil price volatility: comparative analysis of volatility models.

Ng'ang'a, Faith Wacuka
Strathmore Institute of Mathematical Sciences
Strathmore University

Strathmore
UNIVERSITY

MODELLING AND FORECASTING OF CRUDE OIL PRICE VOLATILITY

Comparative Analysis of Volatility Models

NG'ANG'A FAITH WACUKA

Submitted in partial fulfilment of the academic requirement for The Degree of Masters of Science in Mathematical Finance at Strathmore University

Institute of Mathematical Sciences
Strathmore University
Nairobi, Kenya

September, 2021

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Ng'ang'a Faith Wacuka

28^{th} September 2021

Approval

The thesis of Ng'ang'a Faith Wacuka was reviewed and approved by the following:

Mr. Meleah Oleche,

Lecturer, Strathmore Institute of Mathematical Sciences

Strathmore University

Dr. Godfrey Madigu,

Dean, Strathmore Institute of Mathematical Sciences,

Strathmore University

Dr. Bernard Shibwabo,

Director of Graduate Studies,

Strathmore University

Abstract

This study aims at providing an in-depth analysis of forecasting ability of different GARCH models and to find the best GARCH model for VaR estimation for crude oil. The VaR forecasting performance of GARCH-type models are analyzed and compared in a long horizon; based on the Kupiecs POF-test and Christoffersens interval forecast test as well as a Backtesting VaR Loss Function.Crude oil is one of the most important fuel sources and has contributed to over a third of the world's energy consumption. Oil shocks have influence on macroeconomic activities through various ways. Sharp oil price changes delay business investment because they raise uncertainty thus reducing aggregate output for some time. Modelling and forecasting of crude oil prices plays a significant role in supporting policy and decision making in the economy. Successive developments of models used provide opportunities to analyse crude oil market in depth and improve the accuracy of oil price forecasting. The study uses Brent Crude Oil prices data over a period of ten years from the year 2011 to 2020. The study finds that the IGARCH-T distribution model is the best model out of the five models for VaR estimation based on LR.uc Statistic (0.235) and LR.cc Statistic (0.317) which are the least among the values realized. ME and RMSE for the five models used for forecasting have negligible difference. However, the IGARCH model stands out with IGARCH- T-distribution being the best out of the five models in this study with ME of 0.0000963591 and RMSE of 0.05304335. We therefore conclude that the IGARCH- T distribution model is the best model out of the five models used in this study for forecasting Brent crude oil price volatility as well as for VaR estimations.

Contents

Abbreviations.

ARCH – autoregressive integrated moving-average

GARCH - generalized autoregressive conditional heteroskedasticity

EGARCH –exponential GARCH

IGARCH- Integrated GARCH

TGARCH- Threshold GARCH

GJR GARCH- Glosten-Jagannathan-Runkle GARCH

AIC - Akaike information criterion (AIC)

BIC - Bayesian information criterion (BIC)

DF- Dickey Fuller

ADF - Augmented Dickey-Fuller

PP – Philips Peron

GDP – Gross Domestic Product

VaR- Value at Risk

Acknowledgement

I would like to thank Mr. Meleah Oleche, the project supervisor, whose support and help has led to the success of this project. Secondly, I appreciate Prof. Livingstone Luboobi who has walked with me through the process.

I would also like to thank my colleagues in the MSc. Mathematical Finance Class, for their support, friendship and encouragement.

I am grateful to the Strathmore Institute of Mathematical Sciences (SIMS) for their continued support during my time at Strathmore University.

I am also grateful to my family and friends for their support and prayers.

Finally, I thank God Almighty for providing me with the strength and grace through the process.

Dedication

This thesis is dedicated to my family, friends, workmates and classmates for their support and prayers throughout the entire time I was working on the thesis.

Chapter 1

1 Introduction.

1.1 Background.

Commodity markets are briskly expanding and becoming more interesting to many investors in the financial world. The energy market has gained interest continually due to the volatile nature of energy commodity prices. The fundamental assets in energy markets include: electricity, oil, gas and carbon emissions. Past investigations in the field have shown that the prices involved in such markets have four major characteristics: seasonality, spikes, mean reversion and non-storability. The large variations in crude oil prices observed in recent years have caused a lot of concern among participants and regulators in the market. This is because oil price uncertainty has a significant impact on the economy.

Oil price volatility show the magnitude of the rise and fall of oil prices over a period of time. Volatility arises when some error terms are larger than the rest. These error terms become responsible for the unique behaviour of the series. This is known as heteroscedasticity. Past studies have shown that besides global macroeconomic and geopolitical conditions, crude oil prices are also influenced by the fluctuations in supply and demand. Oil is a vital commodity and therefore, controlling the oil market is critical in ensuring that global supply and demand of oil is met. As a result of this need to control the market, there are some organizations created that play an important role in controlling this global supply and demand. One such organization is The Organization of the Petroleum Exporting Countries (OPEC). It aims at coordinating and unifying the petroleum policies of its member countries and ensuring the stabilization of oil markets, so as to secure an efficient, economic and regular supply of petroleum to consumers, steady earnings for producers, and fair returns on capital for those investing in the petroleum industry. The organization also provides information regarding international markets for its member countries.

Forecasting oil price and measuring its risk have been popular topics. The most commonly used measurement for the risk estimation is the Value-at-Risk (VaR for short), which measures the maximum loss of a portfolio value over certain time period by a given confidence level. Although there are some theoretical studies about the shortcoming of VaR due to its lack of sub-additivity and convexity (Cheng, Liu, and Wang (2004)), Orhan & Köksal (2012) acknowledge that there is still no better measure to quantify risk. Identifying proper GARCH-type models with appropriate distributions to evaluate VaR of oil price has become one of most important goals for risk measurement in the crude oil market. In order to improve the measure for VaR, an investor needs to estimate the volatility of crude oil price, i.e. risk.

Research in modeling oil price volatility is inadequate. As seen by Sardorsky (2006) and Narayan (2007), most studies use a particular structure of models to analyse time series and fail to pay attention to use of appropriate model selection criteria to determine what model to use as well as validate the choice of the model. An essential detail to note is that since volatility is time varying, the choice of appropriate model may change over the period of time.

The main objective of this study is to contribute to the development of energy markets through analysis of models used in modelling and forecasting crude oil price volatility and VaR estimation models. The focus is on GARCH family models which have been widely used to model time series. Time series is a set of well-defined data items collected at successive points at uniform time intervals. Time series analysis is an important part of statistics that analyses the data set to study the characteristics of the data and helps in predicting future values of the series based on the characteristics, (Goswami et.al, 2014). Over the last few years, modeling volatility of a financial time series has become a frequently researched area because volatility is considered an important concept for risk management and asset pricing. Volatility from the different GARCH models is then compared to test the significance of each models. The VaR results for the five models are also compared to find out if better VaR results for a GARCH models translates to better forecasting model. This study will; analyse properties of the selected data, ascertain existence of volatility in oil spot price, fit the models, perform VaR backtesting, forecast volatility using selected symmetric and asymmetric models, select the best model by comparing the volatility results from the models and provide an analysis to validate the choice of the selected models.

1.2 Problem Statement

There is no conclusive answer as to which is the best volatility model for forecasting volatility and VaR estimation. Changes in oil prices have a direct effect on prices of other commodities. As investors seek to estimate future prices enabling them to make investment decisions, it is equally important to measure the risk of loss for investments. Elder and Serletis (2010) in their study conclude that when oil price changes are unpredictable, there is a significant drop in real output. An investor's goal is to always obtain maximum returns while taking considerably low risk. The use of GARCH models in VaR estimation and forecasting cannot be disputed. Aloui (2010) computed the VaR using FIGARCH, FIAPARCH and HYGARCH. Smolovic et al (2015) study the performance of eight GARCH models and conclude that the TS GARCH, T GARCH and EGARCH are best for forecasting volatility in the Montenegrin emerging market.

However, there are many contradicting conclusions regarding the choice of volatility models to use. This poses a challenge to investors as the choice of volatility model directly affects the performance of the investor. Vlaar (2000) tested the GARCH model under different distribution assumptions on Dutch bond portfolios and concluded that the GARCH model under the Normal distribution dominates the common practice of using historical simulation models. Miletic and Miletic (2015) showed that GARCH models with a t-distribution of residuals in most analysed cases give a better VaR estimation than GARCH models with normal errors in the case of a 99% confidence level, while the opposite is true in the case of a 95% confidence level. Orhan & Köksal (2012) concluded that the ARCH model and leptokurtic error distributions yielded the best results for VaR estimations while Angelidis et. al (2004) found no clearly superior model but concluded that leptokurtic distributions outperformed the Normal distribution; especially for the ARCH model. Few studies have been done focusing on both VaR estimation and volatility forecasting under the same conditions.

This study sheds light on the issue of volatility forecasting under risk management environment and on the evaluation procedure of various risk models while considering different distributions.

1.3 Objectives

The objectives of this study are:

1.3.1 Main Objective

To perform comparative analysis of different GARCH models to test their ability to forecast prices and estimate value at risk for Brent Crude Oil.

1.3.2 Specific Objectives

- To fit GARCH models and perform backtesting to find the best GARCH model for VaR
- To estimate volatility using different GARCH family models
- To carry out comparative study of different volatility models and validate the choice of the model used for forecasting

1.4 Significance of the study

Proper analysis of crude oil prices trends will help market participants to make informed choices and come up with policies to assist in trading. This will in turn open up the energy market to investors seeking a return, as well as traders who may wish to purchase the commodity. This study therefore aims at contributing to this through modelling the price volatility of crude oil. The focus is on comparative analysis of price volatility models where volatility estimated using different GARCH models is compared.The study also takes into account VaR where VaR results are compared with forecasting results to find out if the best GARCH model for VaR gives the best volatility forecast. The thesis is arranged as follows: In Chapter 2 we review existing literature on energy markets and crude oil price volatility; In Chapter 3 we discuss the models to be used in modelling crude oil prices i.e. GARCH, EGARCH, IGARCH, TGARCH and GJR GARCH and various tests done to pick the best model; In Chapter 4, we discuss the parameter estimation results, VaR results, forecasting results and comparative analysis results. Finally, in Chapter 5, we conclude our study.

Chapter 2

2 Literature Review

This section reviews relevant studies on: Past studies on energy commodities, market volatility and characteristics of oil markets, Causes of crude oil price volatility, Models used in modeling volatility, VaR estimation and Impact of crude oil price volatility.

Characteristics of oil markets

It is important to understand the characteristics of oil markets in order to model the prices. Morana (2001) and Bina & Vo (2007) acknowledge that a large number of empirical studies have concluded that characteristics of time series of crude oil prices are mainly volatility clustering, fat tail distribution, asymmetry and mean reversion. Askari and Krichene, (2008) in their study find that oil price dynamics during this period were characterized by high volatility, intensity jumps and strong upward drift. They also note that the oil price dynamics in the same period were associated with oil markets and world economy's underlying fundamentals.

Doran & Ronn (2007) find that the inclusion of a market price of volatility risk is necessary to capture the degree of bias. They conclude that the market price of volatility risk is negative and significant for natural gas, crude oil, and heating oil and there is a seasonality in the volatility risk premium for natural gas.

Causes of crude oil price volatility

In order to be successful in the oil market, market participants and regulators strive to know how price changes occur as well as how markets react to these shocks. Kilian (2009), acknowledges that it is crucial to know the cause of a given oil price change. Is it demand or supply? This is because the action to take when it comes to stock prices, dividend yield components, and volatility depends on the origin of the oil price shock since they affect variables such as inflation and GDP.

Bastianin & Manera (2015) expound on the work of Kilian and Park (2009). They study the impact of oil price shocks on stock market volatility using Kilian's (2009) structural Vector Autoregressive (VAR) model. Crude oil price changes are modelled as arising from three different sources: i.e. shocks to the supply of crude oil, to the aggregate demand for all industrial commodities and to oil-specific demand. They find that the impact of supply shortfalls is negligible and that volatility responds majorly to shocks hitting the aggregate and oil-specific demand.

Degiannakis et al. (2014) also build on Kilian (2009) work. They use the model by Kilian (2009) and research on the response of volatility to structural oil market shocks. They propose that volatility reacts only to unforeseen changes in total demand. Thus the supply-side and oil-specific demands have no role in market shocks. Demirbas et al (2017) in their study conclude that oil price volatility depends on the combined effects of a number of invariant and variable factors. They also find that a drop in oil prices is brought about more by supply factors than demand factors.

Models used in modeling crude oil price volatility and VaR estimation

From the above studies, it is evident that a lot of researchers have dedicated their time and resources in trying to understand crude oil volatility and risk measurement.

Value-at-Risk (VaR for short),measures the maximum loss of a portfolio value over certain time period by a given confidence level. GARCH models are some of the most commonly used models for VaR estimations. Fan et al.(2008) estimated VaR of crude oil price using GARCH models,based on the Generalized Error Distribution (GED) and detected extreme risk spillover effect between Europe Brent and West Texas Intermediate (WTI) markets. Hung et al.(2008) investigated the influence of fat-tailed process on the performance of one-day-ahead VaR estimates about energy commodities using three GARCH models.

In order to improve the measure for VaR, an investor needs to estimate the volatility of crude oil price. Due to the presence of heteroscedasticity in the variance of financial instruments, ARCH and GARCH models were developed by Engle (1982) and Bollerslev (1986) respectively. From this, many extensions of the GARCH model have been developed to capture the changing volatility due to different factors in the time series. However, there is no definite answer as to which of the models from the GARCH family is the best at forecasting the volatility for all types of financial data. Thus creating the need to have the models restricted to specific data sets.

GARCH models have been widely used in modelling volatility. Khindanova (2004) states that GARCH models have been developed to model the volatility of finance data.

Arouri et. al. (2010) seek to find out if structural breaks and long memory are applicable in modeling and forecasting the conditional volatility of crude oil spot and futures prices using three types of GARCH models- linear GARCH, GARCH with structural breaks and FIGARCH. They conclude that the long memory evidence seen in the in sample period is not strongly supported by the out-of-sample forecasting.

The GARCH model has been found to be very efficient in modeling volatility. Despite these developments, Wang and Wu (2012) acknowledge that the simple GARCH (1,1) model still is still important. This is due to the fact that the model converges much faster to a local maximum in quasi-maximum likelihood estimation and it still gives good forecasting results just as the other multivariate models.

Lama et al (2015) however studies the simple GARCH model and find some weaknesses. He finds symmetric and asymmetric patterns in some time-series making the simple GARCH model inefficient for such time-series because it deals with the magnitude but not the positivity or negativity of the shocks. The simple GARCH model has thus turned out to be relatively inefficient in modeling and forecasting such series creating the need for extension of the GARCH family model so as to take into account the inefficiencies of simple GARCH models. This need was first answered by Nelson (1991) who developed the EGARCH model.

The use of GARCH models is evidently seen with some of the researches done using the models explained as below; Morana (2001) uses the semi-parametric approach which uses the GARCH properties of Brent Market crude oil price volatility. Fong and See (2002) use a Markov regime-switching model allowing for GARCH-dynamics and sudden changes in mean and variance so as to model the conditional volatility of daily returns on crude-oil futures prices. The researchers note that regime-switching model performs better than non-switching models, no matter the evaluation criteria used in forecasting. Vo (2009) also uses a regime switching stochastic volatility model to studies the behaviour of oil prices of WTI market in order to forecast their volatility. He then models the volatility of oil return as a stochastic volatility process where the mean is affected by regime shifts.

Implication and impact of crude oil price volatility

Liu et al (2013) investigate the short- and long-term cross-market uncertainty transmission implied by OVX (Crude oil volatility index) and other volatility indices. They conclude that there are no strong long-run equilibrium relationships among these volatility indices, indirectly verifying the effectiveness of cross-market volatility portfolio strategies. They also acknowledge that there is significant short-term uncertainty transmission between oil and other major markets. Elder and Serletis (2010) in their study conclude that when oil price changes are unpredictable, there is a significant drop in real output. This crucially affects measures of durable consumption and fixed investment in the US. Backus and Crucini (2000) also acknowledge that oil price changes may be responsible for variations in the international terms of trade.

Price volatility effects in the crude oil market is spreading to non-energy commodity markets. With the developments in the oil sector where fossil fuels have been substituted with bio-fuel and increased hedging of oil prices,the link between crude oil market and agriculture markets and metal markets has increased. Ji & Fan (2012) research focuses on oil price volatility effects on non-energy commodity markets. They achieve this by constructing a bivariate EGARCH model with time-varying correlation construction. They find significant spillover effects of crude oil market on non-energy commodity markets.

Chapter 3

3 Research Methodology

This chapter introduces the models used to estimate crude oil price volatility and for VaR estimations. The choice of the modelling approach employed is determined from the characteristics of the historical crude oil spot price data. Energy prices have complex distribution/ properties and as such there is no widely accepted answer to what the best models of price volatility are. The steps that are followed to model the volatility are as below; analyze historical data to find its properties; Test for stationarity of the series and check presence of ARCH effects; The Augmented Dickey-Fuller (ADF) test is used to test for stationarity and Lagrange multiplier test is done to detect the presence of ARCH effects. Explain the five GARCH models used in this study and their estimation procedures for modelling and forecasting crude oil prices. VaR results are then compared with forecasting results to pick the best fitting model.

3.1 Data

Daily Brent Crude oil spot prices for a period of ten years (2011-2020) are used to model the volatility. The choice of this data is because Brent Crude oil has remained supreme in oil market. Crude oil is either traded by themselves or the prices are reflected in other types of crude oil. The availability of the data is also a factor in making this choice.

Data analysis is carried out to find various statistical properties of the data i.e. Mean, variance, skewness, kurtosis, log-returns, squared log-returns, hereroskedasticity. Existence of volatility from the data is also ascertained.

Lux et al (2015) formalize financial returns as;

$r_t = \mu_t + \sigma_t \epsilon_t$

where $r_t = 100 * [ln(p_t) - ln(p_{t-1})], ln(p_t)$ is the log asset price, $\mu_t = E_{t-1}[r_t]$ is the return series conditional mean, σ_t is the volatility process and ϵ_t is standard normal distribution. Defining $X_t = r_t - \mu_t$, the centered returns are given by;

$X_t = \sigma_t \epsilon_t$ under the assumption that μ_t follows an AR(1) process.

The ARCH Lagrangian Multiplier (LM) test proposed by Engle (1982) is used to ascertain the existence of volatility in the data while **Jarque Bera Test** is used to test normality for the logreturns.

Part of the data is used for model building and validation. The rest of the data is used for forecasting. Analysis of the data is done using the **R Statistical Software**

3.2 Mean Equation Selection

Various models are considered in the mean computation. According to Klein (1997), these models originated in the 1920s in the works of Udny Yule, Eugen Slutsky and other researchers. The first known application of the models was that of Yule in 1927 while analysing the time-series behavior of sunspots.

Autoregressive model (AR)

An example of the AR model is the AR(1) model, given by;

$y_t = \alpha_0 + \alpha x_{t-1} + \epsilon_t$

where ϵ_t is a white noise series with mean 0 and variance σ^2

The general formulae for an AR process, AR(p) is;

$y_t = \alpha_0 + \alpha_1 x_{t-1} + ... + \alpha_p x_{t-p} + \epsilon_t$

Moving Average model (MA)

The general form of an MA(1) model is;

$y_t = \mu + \epsilon_t + \theta\epsilon_{t-1}$

where μ is a constant and ϵ_t is a white noise.

The general form of an MA process, MA(q) model is;

$y_t = \mu + \epsilon_t + \theta\epsilon_{t-1} + ... + \theta_q\epsilon_{t-q}$

where q>0

Autoregressive moving average (ARMA (p,q))

An ARMA model is a combination of the AR and the MA processes. The AR part is a representation of the effects of previous observations. The MA part represents effects of previous random shocks (errors).Specified as follows:

$y_t = \sum \alpha_i y_{t-p} + \sum \theta_i \epsilon_{t-q}$

3.3 Modelling ARCH effects

The GARCH models are used in particular because of the following characteristics of financial time series data:

i. **Fat Tails:** A fat tailed distribution is a probability distribution that exhibits large skewness or kurtosis. When the distribution of financial time series is compared with the normal distribution, fatter tails are observed.

ii. **Volatility Clustering:** The second stylized fact is periods of volatility clustering which mean that large and small values in the log-returns tend to occur in clusters. i.e., the large changes tend to be followed by large changes and small changes tend to be followed by small changes. When volatility is high it is likely to remain high and when it is low, it is likely to remain low. Volatility clustering is nothing but accumulation or clustering of information. This feature reflects on the fact that news is clustered over time (Engle, 2004).

iii. **Mean reverting:** The variance of the financial time series data often deviate from the mean. In most cases the reversion is not by a constant factor a property known as heteroscedaticity.

Checking for ARCH effects

The main test for heteroscedaticity is the Lagrange multiplier (LM) test of Engle (1982). This test is equivalent to the F- statistic for testing $\alpha_i = 0 (i = 1, 2, ...m)$ in the linear regression model.

$y_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + ... + \alpha_m y_{t-m}^2 + \epsilon_t$

where t=m+1, m+2, ..., T

ϵ_t denotes the error term is a pre-specified integer and T is the sample size.

The hypothesis is as follows;

$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = ... = \alpha_m = 0$

$H_1 : alpha_1 = \alpha_2 = \alpha_3 = ... = \alpha_m \neq 0$

According to the ARCH LM test the null hypothesis should be rejected when the p-values are below 0.5.

3.4 Stationarity Tests

A common assumption in many time series techniques is that the data is stationary.A stationary process is one where the process tends to come back to the mean i.e. mean reverting. This means a flat series without trends, constant variance and autocorrelation over time and no seasonality. According to Smigel (2021), stationarity is important in time series analysis because a predictable distribution enables forecasting. A series is said to have weak stationarity if mean and variance are constant but autocovariance and autocorrelation are dependent on the lag length and have strong stationarity if all time series characteristics are not dependent on time i.e. they are constant.There are two types of stationarity. Trend stationarity and Difference stationarity. If the process is found to be non-stationary, detrending of differencing is done to make the process stationary.

The GARCH (p,q) process is weakly stationary iff $\sum_{i=1}^{q} + \sum_{j=1}^{p} b_j < 1$

Detrending

A trend stationary series is as below;

$y_t = \mu + \theta_t + \psi(t)\epsilon_t$ where; θ_t is the trend and

$\psi(t)\epsilon_t = \epsilon_t + \theta\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \dots$

And y_t is a function coefficient of time and other factors. Detrending of y_t makes the process stationary i.e.

$y_t - \theta_t = \mu + \psi(L)\epsilon_t$

Differencing

Suppose we have y_t, $y_t - y_{t-1} = \Delta y_t$. The series y_t could be non- stationary but its change δy_t is stationary. i.e.

$y_t = \mu + y_{t-1} + \psi(L)\epsilon_t$.

Since the process is non-stationary, the coefficient of $y_{t-1} = 1$ i.e. unit root

$y_t - y_{t-1} = \mu_\psi(L)\epsilon_t$ makes the process stationary.

Differencing of a trend stationary process and detrending of difference stationary series can also be done.

Failure to have stationary series produces non-sensible regression results. There are two stationary status/ levels of integration of a series i.e. stationarity at levels and difference stationarity series (first level differencing).

There are various tests for stationarity.

a. Dickey Fuller (DF) test

The DF test is accredited to Dickey & Fuller (1979). Given $y_t = \theta y_{t-1} + \epsilon_t$; $\Delta y_t = (\theta - 1)y_{t-1} + \epsilon_t$

The hypothesis is as below;

$(H_0) : \theta = 1 \ (\theta - 1) = 0$

$H_1 : \theta < 1$; $(\theta - 1) < 0$

If H_0 holds, then the series is non-stationary, if H_1 holds, then the series is stationary.

b. Augmented Dickey Fuller (ADF) test

The ADF is an augmented version of the Dickey–Fuller test for a larger and more complicated set of time series models. The ADF statistic used in the test is a negative number. The more negative it is, the stronger the rejection of the hypothesis that there is a unit root at some level of confidence.

Accredited to Said & Dickey (1984), this test relies on the parametric transformation of the model. Let;

$\Delta y_t = \theta^* y_{t-1} + \theta_1^* \Delta y_{t-1} + \theta_t^* \Delta y_{t-2} + \ldots + \theta_p^* \Delta y_{t-p} + \epsilon_t$

Comparing DF test and ADF test, ADF whitens the errors more.

c. Philips-Permon test This test is accredited to Philips Permon (1988). He suggested a non-parametric correction to ADF to account for autocorrelation associated with breaks/ regime shifts in the data.

While ADF test is a parametric model, PP test uses a non-parametric statistical method to cater for the error term's serial correlation. In this study, **the ADF** is used to test for stationarity because it can be used with serial correlation and is more powerful than DF test.

The hypothesis is:

$H_0 : \theta = 1 : Nonstationary$

$H_1 : \theta < 1 : Stationary$

We reject the null hypothesis when the test statistic is less than 0.5 and accept when it is greater than 0.5. A data series is said to be stationary if its mean and variance are constant over time and the value of covariance between two time periods doesn't depend on time.

3.5 Normality Test

Jarque-Bera test is a goodness-of-fit measure of departure from normality, based on the sample kurtosis and skewness. The test is named after Carlos M. Jarque and Anil K. Bera. According to Jarque and Bera (1980), the test statistic JB is defined as:

$$JB = \frac{n}{6}\left(s^2 + \frac{(k-3)^2}{4}\right)$$

where n is the number of observations, s is the sample skewness while k is the sample kurtosis. The statistic JB tests the null hypothesis that the data is from a normal distribution. The null hypothesis is a joint hypothesis of the skewness being zero and the excess kurtosis being zero, since samples from a normal distribution have an expected skewness of zero and an expected excess kurtosis of zero (which is the same as a kurtosis of 3). As the definition of JB shows, any deviation from this increases the JB statistic.

3.6 The models explained

The Box Jenkins' ARIMA methodology had dominated time-series forecasting for long. However, the need for modelling volatile data gave way to development of other models. The ARCH model introduced by Engle (1982) describes the variance of the current error term as a function of the previous period's actual sizes of the error terms. Engle (1982) also finds that the variance is also related to the squares of the previous error terms. The model is commonly employed in modeling financial time series that exhibit time-varying volatility and volatility clustering. The ARCH model is used when the error variance in a time series follows an AR model; if an ARMA model is assumed, the model becomes a GARCH model which was introduced by Bollerslev (1986). Upon analysis of our data, ARMA (2,4) is found to be the best mean model to use among the selected lags. The study uses symmetric and asymmetric models for comparative analysis.

3.6.1 GARCH Model

The ARCH model and its generalized form GARCH models can capture stylized facts such as long and short memory, clustering effects and asymmetric leverage effects.

ARCH models assume that tomorrow's return variance is an equally weighted average of the squared residuals of the last available data. This assumption may not meet the goal of the model as more recent events/data are viewed to be more relevant and should be allocated higher weights. The model also gives zero weights to events that are more than one month old which should not be the case because the events hold some information which could be important/instrumental in forecasting future volatility. The ARCH model estimates the weights of the parameters and allows the data to determine the best weights to use in forecasting the variance. The ARCH(q) model for the series ϵ_t is determined by specifying the conditional distribution of ϵ_t given news/ information available upto time (t-1). Let ψ_{t-1} denote this new. ARCH (q) model for the series ϵ_t is given by;

$\epsilon_t/\psi_{t-1} \sim N(0, v_t)$

$v_t = a_0 + \sum_{i=1}^{q} a_i \epsilon_{t-1}^2$

where $a_0 >, a_t \geq 0$, for all i and $\sum_{i=1}^{q} a_i < 1$ are the requirements to ensure non-negative and finite unconditional variance of stationary ϵ_t series. The weights are equal to unity.

Generalized ARCH (GARCH) model was developed to overcome ARCH models challenges. Just like ARCH model, GARCH model is a weighted average of past squared residuals. However, its weights reduce but never become zero. The model argues that the best predictor of next period's variance is a weighted average of the long run average variance, the variance predicted for this period and the new information in this period i.e. the most recent squared residual. The model has the form;

$$\epsilon_t = \xi v_t^{1/2}$$

$$v_t = a_0 + \sum_{i=1}^{q} a_i \epsilon_{t-1}^2 + \sum_{j=1}^{p} b_j v_{t-j} \tag{1}$$

where $\xi\ IID(0,1)$. For the conditional variance to be positive is; $a_0 > 0, a_i \geq 0, i = 1,2,3,...q.b_j \geq 0, j = 1,2,...,p$

is sufficient. Let $\eta_t = \epsilon_t^2 - h_t$, GARCH model can be expressed in terms of ARMA model as below;

$$\epsilon_t^2 = a_0 + \sum_{i=1}^{Max(p,q)} (a_i + b_i)\epsilon_{t-i}^2 + \eta_t + \sum_{j=1}^{p} b_j \eta_{t-j} \tag{2}$$

A GARCH model is therefore an extension of ARMA approach to squared series ϵ_t^2. Using ARMA model as our mean model;

$$E(\epsilon_t^2) = \frac{a_0}{1 - \sum_{i=1}^{Max(p,q)} (a_i + b_i)}$$

3.6.2 EGARCH Model

GARCH models assume that only the magnitude and not the positivity or negativity of unanticipated excess returns determine the future volatility. Accredited to Nelson & Cao (1991), the EGARCH model was developed to cater for the shortcomings of the GARCH models. The model allows for asymmetric effects between positive and negative shocks and has no restrictions on the parameters. The conditional variance h_t is an asymmetric function of lagged distributions. In financial time-series, it has been stated that volatility behaves differently depending on if a positive or negative shock occurs(leverage effect) and describes how a negative shock causes volatility to rise more than, if a positive shock with the same magnitude had occurred. The model is as below;

$$\epsilon_t \equiv \xi_t v_t^{1/2}$$

$$ln(v_t) \equiv a_0 + \frac{1 + b_1 B + ... + b_{q-1} B^{q-1}}{1 - a_1 B + ... + a_p B^p} g(\epsilon_{t-1}) \tag{3}$$

where:

$$g(\epsilon_t) \equiv \begin{cases} (\theta + \gamma)\epsilon_t - \gamma E(|\epsilon_t|), & if \epsilon_t \geq 0, \\ (\theta - \gamma)\epsilon_t - \gamma E(|\epsilon_t|), & if \epsilon_t < 0, \end{cases}$$

B is the lag operator such that;

$B_g(\epsilon_t) \equiv g(\epsilon_{t-1})$

Lama et.al (2015) specifies the logarithm of conditional variance and states that the model can also be represented as;

$$ln(v_t) = a_0 + \beta ln(v_{t-1}) + \alpha \left| \frac{\epsilon_{t-1}}{\sqrt{v_{t-1}}} \right| + \gamma \frac{\epsilon_{t-1}}{\sqrt{v_{t-1}}} \tag{4}$$

From the equation above, we can deduce that the leverage effect is exponential not quadratic. The forecasts of the conditional variance are also seen to be non-negative.

3.6.3 IGARCH Model

For the GARCH model, GARCH (p,q) process is weakly stationary iff $\sum_{i=1}^{q} a + \sum_{j=1}^{p} b_j < 1$. Stationarity however does not require such a stringent restriction. The unconditional variance does not depend on time. The covariance is not stationary. Thus; $\sum_{i=1}^{q} a + \sum_{j=1}^{p} b_j \equiv 1$

Developed by Engel and Bollerslev (1986), the model takes the form;

$h_t \equiv a_0 + (1 - b_1)\epsilon_{t-1}^2 + b_2\sigma_{t-1}^2$

The model can also be viewed as a product of omitted structural breaks rather than the result of true IGARCH behaviour

3.6.4 Threshold Garch model (TGARCH)

TGARCH model proposed by Zakoian (1994) is another model that can be used to handle leverage effects. It was designed to divide the distribution of the innovations into disjoint intervals and then approximate a piece-wise linear function for the conditional standard deviation and the conditional variance. If there are only two intervals, the division is normally at zero, i.e., the influence of positive and negative innovations on the volatility is differentiated. The TGARCH model of order can be written as;

$y_t = \beta_0 + \beta_1 X_t + \epsilon_t$

$\sigma_t^2 = \omega + \sum \beta_j \sigma_{t-j}^2 + \sum \alpha_i \epsilon_{t-i}^2 + \sum \gamma_k \epsilon_{t-k} I_{t-k}$

where;

$I_{t-k} = 1$ if $\epsilon_{t-k} < 0$

$I_{t-k} = 0$ otherwise

If $\epsilon_{t-k} > 0$, then it indicates good news. If $\epsilon_{t-k} \leq 0$, it represent bad news.

If $\epsilon_{t-k} \leq 0$, then $I_{t-k} = 1$

Therefore;

$\sum \gamma_k \epsilon_{t-k} I_{t-k} = \sum \gamma \epsilon_{t-k}$ where;

If $\gamma_k > 0$, then bad news increase volatility

If $\gamma_k < 0$, then bad news reduce volatility.

If $\epsilon_{t=k} > 0$ then, $I_{t-k} = 0$

Given the information above, impact on volatility can be summarised as below;

For good news= α_i

For bad news= $\alpha_i + \gamma_k$

3.6.5 GJR GARCH Model

Glosten, Jagannathan and Runkle (1993) proposed a modification of the original GARCH model using a dummy variable to capture asymmetric effects in financial time series. The following is the general formula of the model:

$y_t = \sigma_t \epsilon_t where \epsilon_t \ iid(0, 1)$

$\sigma^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_t^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{p} \gamma_i I_{t-i} \alpha_{t-i}^2$

$$I_{t-i} \equiv \begin{cases} 1 & if \alpha_{t-1} < 0 \\ 0 & if \alpha_{t-i} \geq 0, \end{cases}$$

I_{t-i} is a dummy variable that has a value of 1 when the yield is negative and 0 otherwise. The impact of the news, or shocks to the yield on the volatility depends on the sign of the parameter estimated by this dummy variable. In any case, this model, which allows for different responses of the volatility to positive or negative shocks and supposes that the minimum volatility is observed when there, is no news.

3.7 Backtesting Var (Value at Risk)

VaR can be viewed as a gauge that summarizes the worst loss over a target horizon that will not be exceeded with a given level of confidence (Jorion, 2007). More formally, aR is expressed as;

$$Pr(L > VaR) = \alpha$$

where L is the loss on a given day and α is the significance level. VaR is therefore a quantile in the distribution of profit and loss that is expected to be exceeded only with a certain probability, formally expressed as;

$$p = \int_{\infty}^{-VaR(p)} f_q(X)dX$$

In this thesis; the VaR figures are given using a 1% significance level. VaR is computed using the conditional volatility of returns multiplied by the quantile of a given probability distribution. In our study, normal distribution is assumed thus the quantile is -2.33.
The VaR estimates in this thesis will be evaluated using two tests: an unconditional and a conditional test of coverage originally developed by Kupiec (1995) and Christoffersen (1998) respectively.

Kupiec's test

Kupiec's test was developed to test whether the empirical proportion of violations congregate with the nominal proportion specified by the VaR significance level. According to Kupiec (1995), a likelihood ratio test can be constructed as below;

$$LR_{uc} = 2ln\left[(1 - \tfrac{F}{T})^{T-F}(\tfrac{F}{T})F\right] - 2ln(1-p)^{T-F}p^F$$

where T is the number of out-of-sample estimates and F the observed number of violations. Hence, $\frac{F}{T}$ is the empirical VaR size which follows the binominal distribution so $F\ B(T,p)$. LR_{uc} follows the chi-square distribution with one degree of freedom, i.e. $LR_{uc}\ \chi^2_{(1)}$ under the null hypothesis; $\frac{F}{T} = p$. rejection of the null hypothesis implies that the empirical VaR size is significantly different from the stated VaR significance level, i.e. the nominal size.

Christoffersen's test of independence

Ideally, a violation today does not reveal any information about the likelihood of a violation tomorrow, i.e. the violations occur independently of each other. A disadvantage with Kupiec's test is its ability detect whether the violations occur independently or clustered in a sequence. Christoffersen (1998) developed a test to detect clusters of violations. The advantage with the Christoffersen test of independence is its deference to the conditionality in the volatility forecasts. Good volatility forecasts ought to respond to periods of high and low volatility and subsequently adjust its predictions accordingly after the volatility clusters.

The probability of two subsequent violations are therefore defined as;

$$p_{ij} = P(\eta_t = i|\eta_{t-1} = j)$$

Independence of violations is therefore defined as violations that do not occur in two subsequent days. A drawback with this test is arguably the definition of independence as a violation today followed by a violation the day after tomorrow is not detected in this test.Christoffersen (1998) ,suggests as likelihood ratio test of conditional coverage as shown below;

$$LR_{ind} = -2ln[(1-p)^{T-F}p^F] + 2ln(1-\pi_{01})^{\eta_0 0}\pi_{01}^{\eta_{01}}(1-\pi_1 1)^{\eta_{10}}\pi_{11}^{\eta_{11}}$$

where η_{ij} is the number of observations with the value i followed by j for i,j= 0,1 and

$$\pi_{ij} = \frac{\eta_{ij}}{\sum_j \eta_{ij}}$$

are the corresponding probabilities. $LR_{ind}\ \chi^2_{(i)}$ under the null hypothesis which states that the violations are independently distributed. Hence, a rejection of the null hypothesis infers that the violations are clustered and consequently not independent.

3.8 Model selection criteria

This section explains the model selection criteria used to select the model combination to use.

The Akaike Information Criterion (AIC) introduced by Hirotugu Akaike (1973) is used to select the mean model to be used. AIC, which is a penalized-likelihood criteria ,is an estimate of a constant plus the relative distance between the unknown true likelihood function of the data and the fitted likelihood function of the model. A lower AIC means a model is considered to be closer to the true model.

The loglikelihood is used to select the best model for VaR estimation. The higher the log-likelihood, the better the model. Besides this, the information criteria are also used to pick the model. A good model had the highest loglikelihood or the lowest information criteria. Therefore, a higher log likelihood translates to a low information criteria. The information criteria used in this study are the Akaike, Bayes, Shibata and Hannan-Quinn.

Forecasting performance of the five models is analysed by comparing the errors i.e. comparing the forecasted returns with realized returns. This is done by comparing the mean error (ME), mean absolute error (MAE) and root mean square error (RMSE). The lesser the errors the better the more accurate the model is in forecasting correct return for Brent Crude Oil.

$ME = \frac{1}{n}\sum_{j=1}^{n}(y_i - y_j^*)$

$MAE = \frac{1}{n}\sum_{j=1}^{n}|y_i - y_j^*|$

$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_i - y_j^*)^2}$

where; y^* are the forecasted values and y the realized values

Chapter 4

4 Research Findings and Discussion

This study involves analysis the volatility of Brent Crude oil spot prices and VaR estimation using five models. Brent crude oil spot prices from the year 2011 to 2020 were used in the analysis. Part of the data was used for model building (in-samples)and the rest for out of sample forecasting.

4.1 Data Exploration

Spot Prices

The analysis is done by use of time series plots and descriptive analysis .The plot for Brent Crude Oil Prices for the period between 2011 and 2020 is as in figure 1

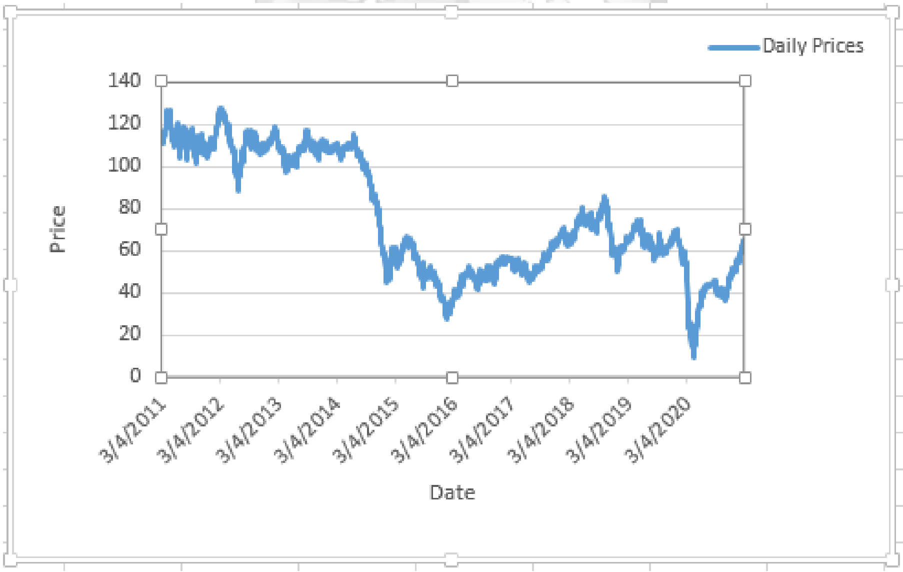

Figure 1: Daily Brent Crude Oil closing price

From figure 1. it can be seen that the spot prices are highly volatile and have had a downward trend from the year 2011 to 2020. It is important to note the high drop in oil prices earlier in the year 2020 which was fueled by the Covid-19 pandemic. This is a good example of how oil prices are affected by other factors in the world.Volatility clustering is also evident from the table above i.e. prices will rice continuously for a period of time and drop continuously for a period of time

Figure 2. below is a qqplot for the spot prices which shows that the spot prices for the period are not normally distributed.

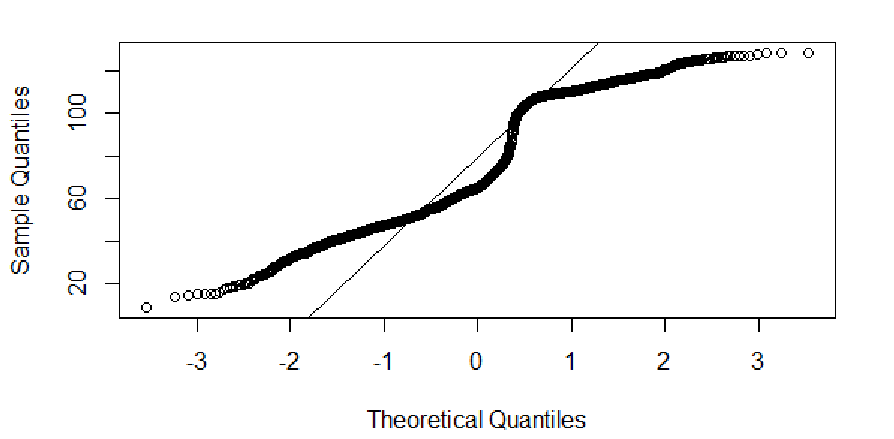

Figure 2: qqplot for Brent Crude Oil spot prices

Returns

To analyse the volatility of crude oil prices, we use oil returns for the same period. Returns can be comparable with each other while prices on the other hand always depend on the previous price. Returns are preferred when modelling volatility because prices are bounded to be non-negative and usually have a unit root, while log-returns can have any value, which makes them easier to model and one can achieve stationarity using log returns.

Time plots are used to determine the observable characteristics of the returns as presented in Figure 3 and 4 which show the log returns and squared logreturns respectively.

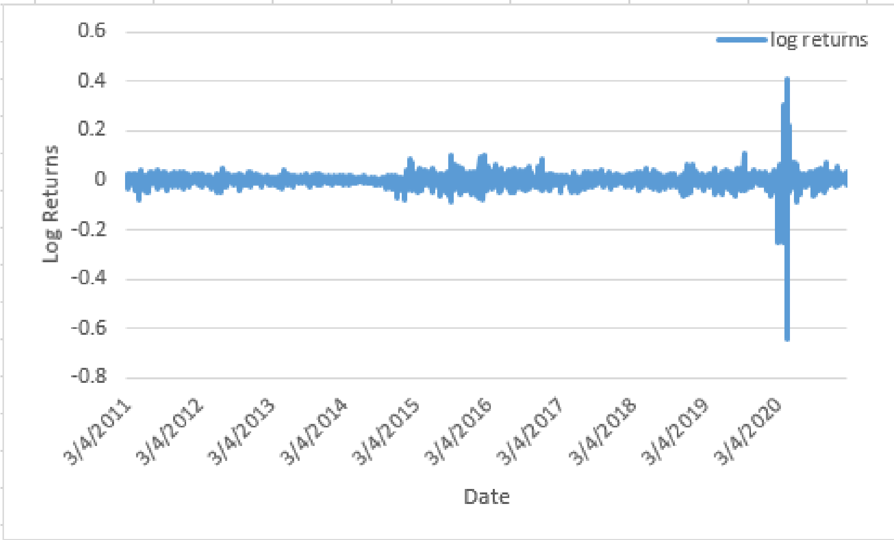

Figure 3: Daily Brent Crude Oil Log Returns

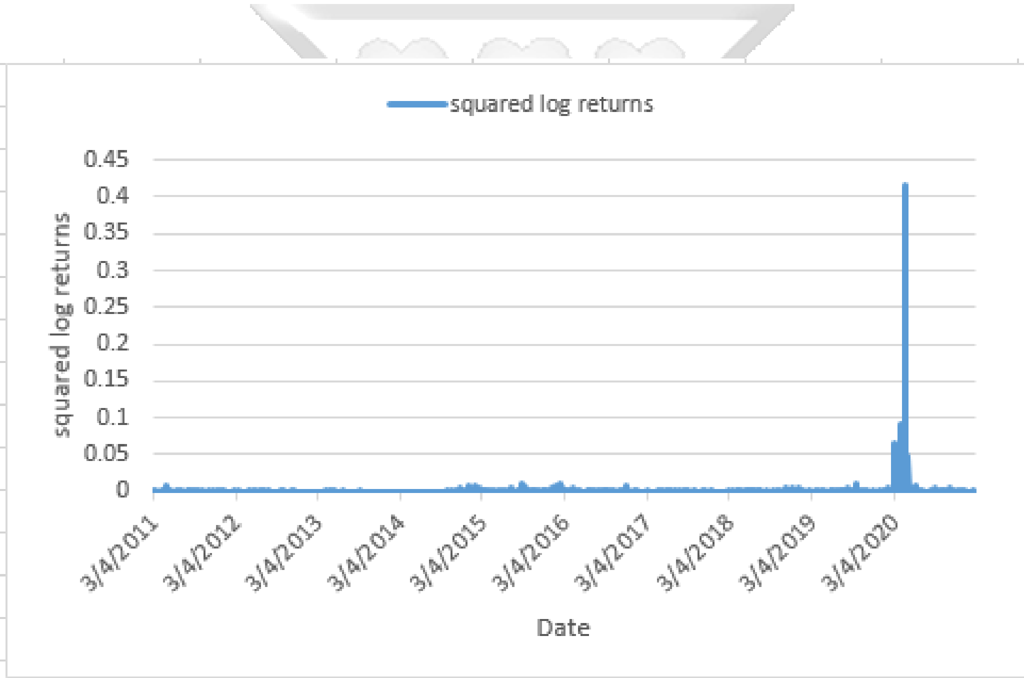

Figure 4: Daily Brent Crude Oil squared log returns

From the plots it is evident that the financial time series exhibit common features. Variance is not constant throughout the figures, which is evidence of heteroscedasticity/ mean reversion property. Volatility clustering can also be seen from the plots.

The ARCH Langrangian Multiplier Test further confirms these findings that there is presence of heretoscedasticity in the crude oil data.

```
aarch.test(logreturns,arch="box",alpha=0.05,lag.max = 2)
```

```
Box-Ljung test


data:  y^2
X-squared = 367.14, df = 2, p-value < 2.2e-16
alternative hypothesis: y is heteroscedastic
```

The descriptive statistics which includes; mean, standard deviation, kurtosis, skewness are utilized so as to describe the returns characteristics as shown in table 1.

Table 1: Returns characteristics.

sample size	2530
mean	-0.0002295898
standard deviation	0.02912715
skewness	-3.324808
kurtosis	126.2278

The kurtosis is greater than three and indicates that the returns have fat tails heavier than a normal distribution. The skewness is not equal to zero which indicates that the returns are not symmetric.

The figure 5 further helps us understand how the returns compare with the normal distribution of the same mean and standard deviation. Brent crude oil log returns plot have the same shape as normal returns curve but have a steaper curve than that of a normal distribution This shows presence of heavier tails than normal distribution.

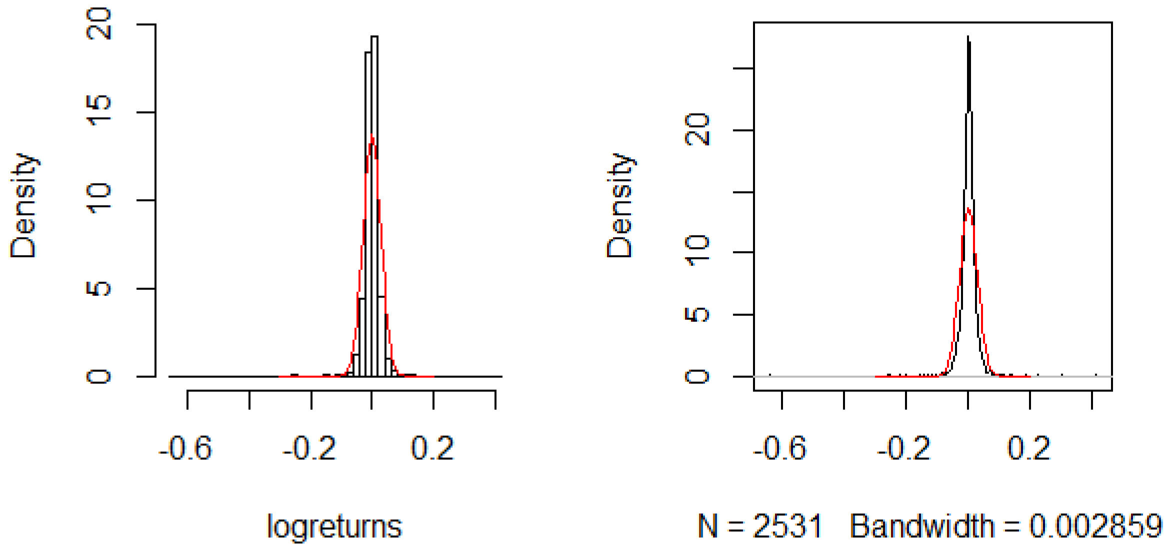

Figure 5: Comparison with normal distribution

4.2 Testing for Normality

The normal QQ-plot is used to analyze the distributional properties, that is, to check whether the return series is normally distributed. The normal QQ-plot represents a scatter plot of a given distribution. The greater the departure from this line the greater the evidence against the null hypothesis of being a normal distribution.

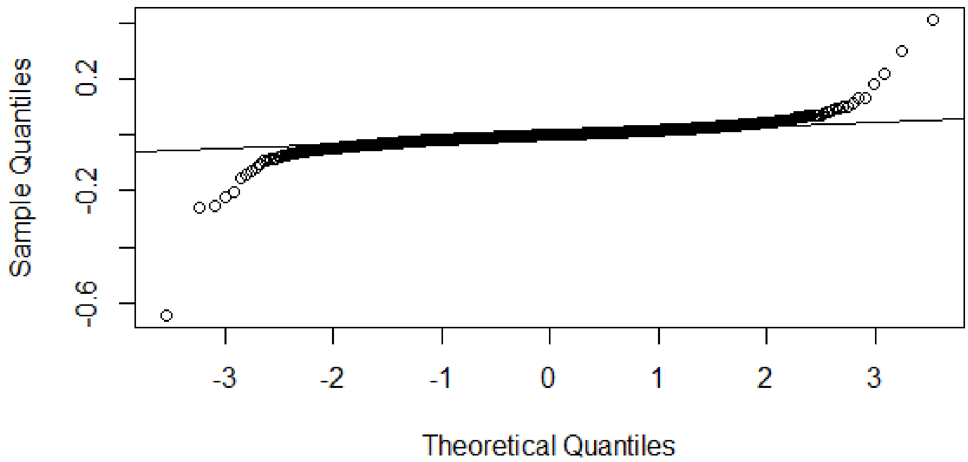

Figure 6: Returns QQ-plot

The plot shows that the returns are relatively normally distributed with some outliers (appear further from the normal line) which can be taken to be the heavier tails seen in figure 5 above. This plot show that normal distribution can be used to model the returns in this study but would not take care of the heavy tails.

The presence of heavy tails prompts us to consider the student t distribution which is known to have the ability to capture heavy tails. The plot below shows QQ-plot for Brent Crude Oil prices fitted using t-distribution.

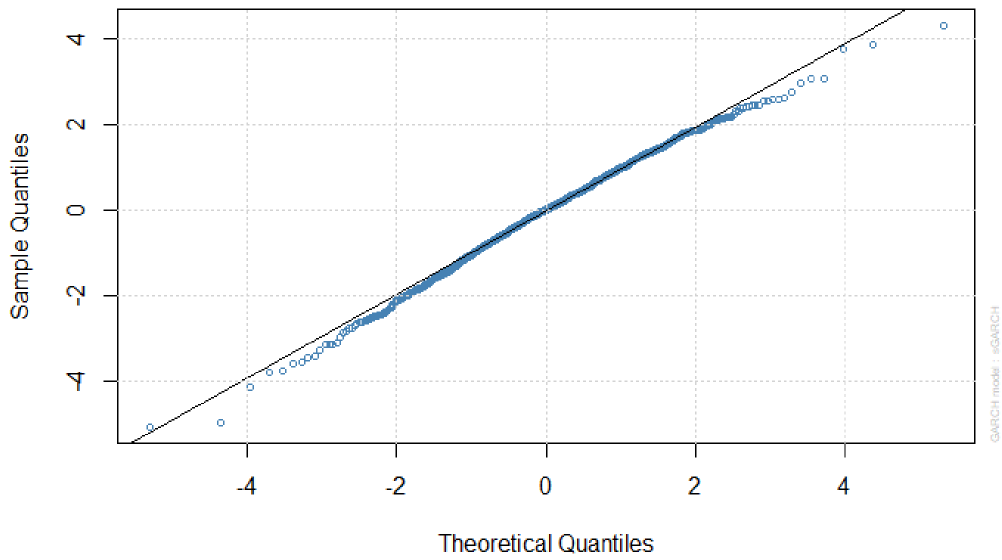

Figure 7: Returns student t QQ-plot

The Jarque Bera (JB) test is applied so as to strengthen these results. The hypothesis is:

H_0 The data is normally distributed

H_1 The data is not normally distributed

```
jarque.bera.test(logreturns)


Jarque Bera Test


data:  logreturns
X-squared = 1607358, df = 2, p-value < 2.2e-16
```

The unusually high value of the Jarque-Bera statistics shows that the null hypothesis of normality can be rejected at the 1of significance, also as evidenced by a high excess kurtosis and negative skewness. The student t distribution is symmetric and bell-shaped, like the normal distribution.

However, the t-distribution has heavier tails thus more prone to producing values that fall far from its mean. Since we cannot fully discredit normal distribution in this study, we chose to fit the volatility models using the normal distribution and t-distribution. T-distribution was added due to its ability to capture heavier tails which are evident from our analysis above.

4.3 Testing for Stationarity

To investigate whether the return series are stationary, the Augmented Dickey-Fuller (ADF) test was applied. The hypothesis was such that:

H_0 :Non stationary

H_1 : Stationary

The results for the test are as shown in Table 2

Table 2: ADF Test

Dickey-Fuller	-11.324
Lag order	13
p-value	0.01

The p-value is <0.05. This allows the rejection of the null hypothesis.

To confirm the results above, we use the Philips Perron (PP) test which is a non-parametric correction to ADF to account for autocorrelation associated with breaks/ regime shifts in the data.

The results for the PP test are as shown in Table 3.

Table 3: PP Test

Dickey-Fuller	-52.403
Truncation lag parameter	8
p-value	0.01

The p value for the PP test is less than 0.05 and is same as the p value for the ADF test. We therefore conclude that that the return series is stationary.

4.4 Determining the mean equation

The first step in the selection of the mean equation is to plot the sample Autocorrelation function (ACF) and Partial autocorrelation function (PACF). If the sample ACF dies off and the PACF cuts off at lag p that would indicate an AR (p) process. Similarly if the sample PACF dies off and the ACF cuts off at lag q the process would be an MA (q) process.

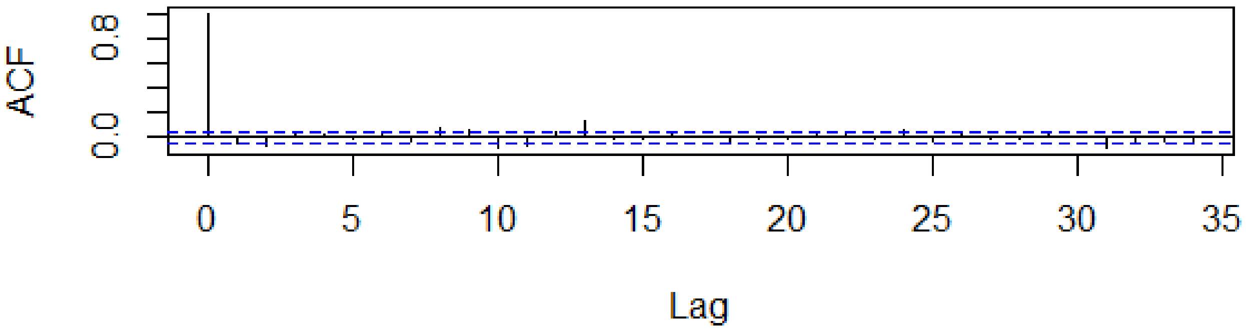

Figure 8: Returns ACF

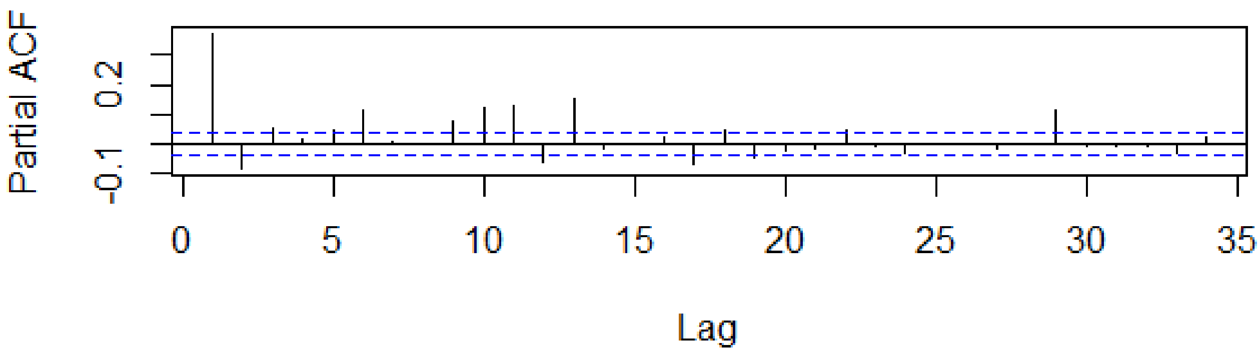

Figure 9: Series squared logreturns PACF

Table 4 shows the AIC values for the selection of the mean equation.

From Table 4, ARMA (2,0,4) has the least AIC and is thus selected as the mean equation. This is in line with the finding that our series is stationary. No differencing or detrending was done to make the series stationary. Therefore, we will use ARMA mean equation and not ARIMA.

Table 4: Mean Model Selection

Models	Order	AIC	Models	Order	AIC
AR	4	-10,307.51	MA	4	-10,307.46
AR	5	-10,305.60	MA	5	-10,305.49
ARMA	(4,3)	-10,307.95	ARIMA	(4,1,3)	-10,294.97
ARMA	(4,4)	-10,306.46	ARIMA	(4,1,4)	-10,288.58
ARMA	(4,2)	-10,303.52	ARIMA	(4,1,2)	-10,291.86
ARMA	(3,3)	-10,309.97	ARIMA	(2,1,3)	-10,298.55
ARMA	(3,4)	-10,307.99	ARIMA	(5,1,3)	-10,288.80
ARMA	(2,4)	-10,319.42	ARMA	(2,3)	-10,310.00
ARMA	(5,3)	-10,305.98	ARMA	(0,0)	-10,312.58
ARMA	(5,5)	-10,302.66	.	.	.

4.5 Modelling and forecasting volatility

This section aims at fitting the models that are being used to compare the volatility. This is done by;

1. Checking for the ARCH effects
2. Fitting volatility models
3. Back-testing
4. Forecasting

4.5.1 Testing for ARCH effects

The Lagrange multiplier (LM) test by ENGLE (1982) was applied to the residuals of simple time series models. The hypothesis is as follows:

$H_0 : \alpha_1 = \alpha_2 = ... = \alpha_m = 0$

$H_1 : \alpha_1 = \alpha_2 = ... = \alpha_m \neq 0$

The ARCH-LM tests results provide strong evidence for rejecting the null hypothesis as shown in table 4 below. This confirm that ARCH effects exist; hence, an ARCH or GARCH model should be employed in modeling the return time series.

Table 5: ARCH-LM Test

.	Brent
Chi-squared	233.24
df	10
p-value	$2.2\epsilon^{-16}$

4.5.2 Fitting volatility models

Although it is rather difficult to estimate the order (p, q), some studies have found that the predictive effect of higher order model is not necessarily better than the low order model (Hansen, P. R., Lunde, A.,2005 and Bollerslev, T., Chou, R.Y., Kroner, K.F 1992). Therefore, we choose (p, q) = (1, 1) for various GARCH models in this study. GARCH, EGARCH, IGARCH, TGARCH AND GJR GARCH models were fitted.

The parameters for GARCH normal distribution are as shown in Table 6;

Table 6: Models' parameters - Normal distribution

.	GARCH	.	EGARCH	.	IGARCH	.	TGARCH	.	GJR
μ	-0.000251	μ	1.04682	μ	-0.000255	μ	-0.000720	μ	-0.000576
ar1	-0.323653	ar1	0.11040	ar1	-0.323667	ar1	-0.351039	ar1	-0.350779
ar2	-0.987638	ar2	0.10861	ar2	-0.987639	ar2	-0.989692	ar2	-0.991159
ma1	0.349572	ma1	0.40827	ma1	0.349564	ma1	0.369471	ma1	0.373420
ma2	1.017209	ma2	0.05510	ma2	1.017128	ma2	1.006488	ma2	1.007623
ma3	0.029357	ma3	0.50047	ma3	0.029320	ma3	0.023553	ma3	0.026635
ma4	0.008236	ma4	0.03617	ma4	0.008168	ma4	0.001602	ma4	0.000316
omega	0.000001	omega	0.66694	omega	0.000001	omega	0.000072	omega	0.000001
alpha1	0.055658	alpha1	0.29317	alpha1	0.057194	alpha1	0.039121	alpha1	0.000000
beta1	0.942370	beta1	0.73588	beta1	0.942806	beta1	0.966937	beta1	0.967021
.	.	$\gamma1$	0.13412	$\gamma1$	.	etal1	0.815349	$\gamma1$	0.063959

The parameter for GARCH- student t distribution are as shown in Table 7;

The loglikelihood and information criteria values for the five models fitted with normal distribution are as shown in Table 8;

The loglikelihood and information criteria values for the five models fitted with student t distribution are as shown in Table 9;

Table 7: Models' parameters - student's t distribution

.	GARCH	.	EGARCH	.	IGARCH	.	TGARCH	.	GJR
μ	-0.000074	μ	-0.000378	μ	-0.000074	μ	-0.000322	μ	-0.000287
ar1	-1.152015	ar1	-1.475760	ar1	-1.149651	ar1	-1.475475	ar1	-1.414132
ar2	-0.328867	ar2	-0.990453	ar2	-0.327741	ar2	-0.990406	ar2	-0.935428
ma1	1.151574	ma1	1.475957	ma1	1.149044	ma1	1.473940	ma1	1.414838
ma2	0.343494	ma2	1.002337	ma2	0.342293	ma2	1.000491	ma2	0.946654
ma3	0.029833	ma3	0.016494	ma3	0.029902	ma3	0.016033	ma3	0.021115
ma4	0.024049	ma4	0.014478	ma4	0.024163	ma4	0.015380	ma4	0.021447
omega	0.000001	omega	-0.023329	omega	0.000001	omega	0.000053	omega	0.000001
alpha1	0.048755	alpha1	-0.055690	alpha1	0.050463	alpha1	0.036867	alpha1	0.003159
beta1	0.948916	beta1	0.997140	beta1	0.949537	beta1	0.969347	beta1	0.968482
.	.	$\gamma 1$	0.064968	$\gamma 1$	.	etal1	0.775855	$\gamma 1$	0.053914
shape	6.546236	shape	7.645430	shape	6.337109	shape	7.255139	shape	7.638457

Table 8: Loglikelihood and Information Criteria values - normal dist

.	GARCH	EGARCH	IGARCH	TGARCH	GJR GARCH
loglikelihood	5398.784	5423.502	5398.583	5414.006	5422.328
Akaike	-5.3249	-5.3483	-5.3257	-5.3389	-5.3472
Bayes	-5.2972	-5.3178	-5.3007	-5.3084	-5.3167
Shibata	-5.3249	-5.3484	-5.3257	-5.3390	-5.3472
Hannan-Quinn	-5.3147	-5.3371	-5.3165	-5.3277	-5.3360

A model that has the highest logarithm maximum likelihood function value is picked as the best model. The less the information criteria values, the better the model. From the two tables above, The EGARCH model has the highest log likelihood value and the lowest Information criteria values for both normal distribution and t-distribution. However, the t-distribution fits the data better as it has higher log likelihood value and less information criteria compared to normal distribution. This show that the EGARCH-t distribution is the best fitted model among the models tested.

News impact curves, introduced by Pagan and Schwert (1990) and Engle and Ng (1991), are useful tools to visualize the magnitude of volatility changes in response to shocks. The name

Table 9: Loglikelihood and Information Criteria values -student's t dist

.	GARCH	EGARCH	IGARCH	TGARCH	GJR GARCH
loglikelihood	5430.455	5446.416	5430.239	5443.052	5442.635
Akaike	-5.3552	-5.3700	-5.3560	-5.3667	-5.3662
Bayes	-5.3247	-5.3367	-5.3282	-5.3334	-5.3330
Shibata	-5.3553	-5.3700	-5.3560	-5.3667	-5.3663
Hannan-Quinn	-5.3440	-5.3578	-5.3458	-5.3544	-5.3540

comes from the interpretation of shocks as news influencing the market movements. They plot the change in conditional volatility against shocks in different sizes, and can concisely express the asymmetric effects in volatility.

Below are news impact curves for some the models;

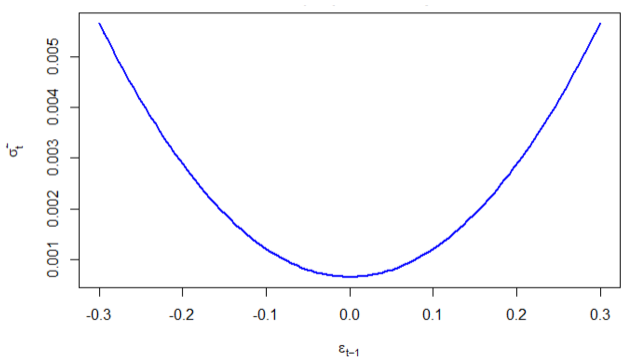

Figure 10: GARCH

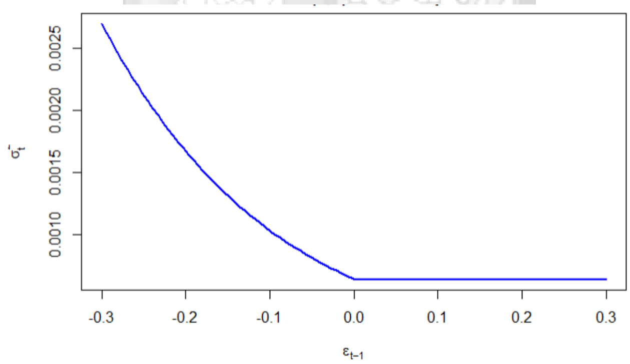

Figure 11: EGARCH

Having fitted the models, we now perform back-testing tests to see whether the models were fitted correctly.

4.6 VaR Backtesting

A popular model to estimate the VaR of financial series is to calculate the number of VaR exceptions, namely days when actual losses exceed VaR predictive results. If the ratio of exceptions is lower than the selected confidence level, the risk is overestimated. On the other hand, too many exceptions implies that risk is underestimated. It is important to note that the exact exceptions suggested by the confidence level is rarely observed. Therefore a statistical analysis is necessary to study whether exceptions is reasonable or not. VaR is usually calculated at the 99% or 95% confidence level. This is the loss that is expected to be exceeded only 1% or 5% of the time respectively.

In this study, we do a historical backtest in checking the model performance where estimated VaR is compared with the actual return over the period. If the return is more negative than the VaR, we have a VaR exceedance. In our case, a VaR exceedance should only occur in 1% of the cases (since we specified a 99% confidence level).

The unconditional coverage test critical value is 3.841459; and the conditional coverage test critical value is 5.991465.

Table 10: Kupiec test - Normal Dist

.	GARCH	EGARCH	IGARCH	TGARCH	GJR GARCH
alpha	1%	1%	1%	1%	1%
Backtest Length	224	224	224	224	224
Expected Exceed	2.2	2.2	2.2	2.2	2.2
Actual VaR Exceed	5	7	5	7	5
Actual %	2.2 %	3.1%	2.2%	3.1%	2.2%
Null-Hypothesis:	.	.	.	.	.
LR.uc Statistic:	2.544	6.535	2.544	6.535	2.544
LR.uc Critical:	6.635	6.635	6.635	6.635	6.635
LR.uc p-value:	0.111	0.011	0.111	0.011	0.111
Reject Null:	NO	NO	NO	NO	NO

For Christoffersen test, the null hypothesis is: Correct Exceedances and Independence of Failures

Table 11: Christoffersen test - Normal dist

.	GARCH	EGARCH	IGARCH	TGARCH	GJR GARCH
alpha	1%	1%	1%	1%	1%
Backtest Length	224	224	224	224	224
Expected Exceed	2.2	2.2	2.2	2.2	2.2
Actual VaR Exceed	5	7	5	7	5
Actual %	2.2 %	3.1%	2.2%	3.1%	2.2%
Null-Hypothesis:	.	.	.	.	.
LR.cc Statistic:	2.773	6.989	2.773	6.989	2.773
LR.cc Critical:	9.21	9.21	9.21	9.21	9.21
LR.cc p-value:	0.25	0.03	0.25	0.03	0.25
Reject Null:	NO	NO	NO	NO	NO

Table 12: Kupiec test- T dist

.	GARCH	EGARCH	IGARCH	TGARCH	GJR GARCH
alpha	1%	1%	1%	1%	1%
Backtest Length	224	224	224	224	224
Expected Exceed	2.2	2.2	2.2	2.2	2.2
Actual VaR Exceed	4	5	3	5	4
Actual %	1.8 %	2.2%	1.3%	2.2%	1.8%
Null-Hypothesis:	.	.	.	.	.
LR.uc Statistic:	1.133	2.544	0.235	2.544	1.133
LR.uc Critical:	6.635	6.635	6.635	6.635	6.635
LR.uc p-value:	0.287	0.111	0.628	0.111	0.287
Reject Null:	NO	NO	NO	NO	NO

Table 13: Christoffersen test- T dist

.	GARCH	EGARCH	IGARCH	TGARCH	GJR GARCH
alpha	1%	1%	1%	1%	1%
Backtest Length	224	224	224	224	224
Expected Exceed	2.2	2.2	2.2	2.2	2.2
Actual VaR Exceed	4	5	3	5	4
Actual %	1.8 %	2.2%	1.3%	2.2%	1.8%
Null-Hypothesis:	.	.	.	.	.
LR.cc Statistic:	1.279	2.773	0.317	2.773	1.279
LR.cc Critical:	9.21	9.21	9.21	9.21	9.21
LR.cc p-value:	0.528	0.25	0.853	0.25	0.528
Reject Null:	NO	NO	NO	NO	NO

From Tables 10 to 13, it can be informed that both unconditional and conditional L.R. statistics are smaller than critical value, which show that both Kupiec test and Christoffersen test don't reject null hypothesis on 1%.These results also indicate that the models can produce accurate VaR forecasts and handle the ever-changing fluctuations in the return rate of Brent Crude Oil. The backtesting results for the five models are summarised as in Tables 10,11,12 and 13. The backtesting results for the five models for normal distribution and t-distribution are summarised as in table 7 and 8. All GARCH-class models pass both LRuc and LRcc tests. The simple GARCH, IGARCH and GJR models outperform the rest fof the normal dostribution. The IGARCH model is the best for t-distribution with the GARCH and GJR models following closely. The LR.uc Statistic and LR.cc Statistic for the IGARCH- T distribution are the least. We therefore conclude that the IGARCH- T distribution model is the best for VaR estimations.

The plots showing the backtesting results are as below;

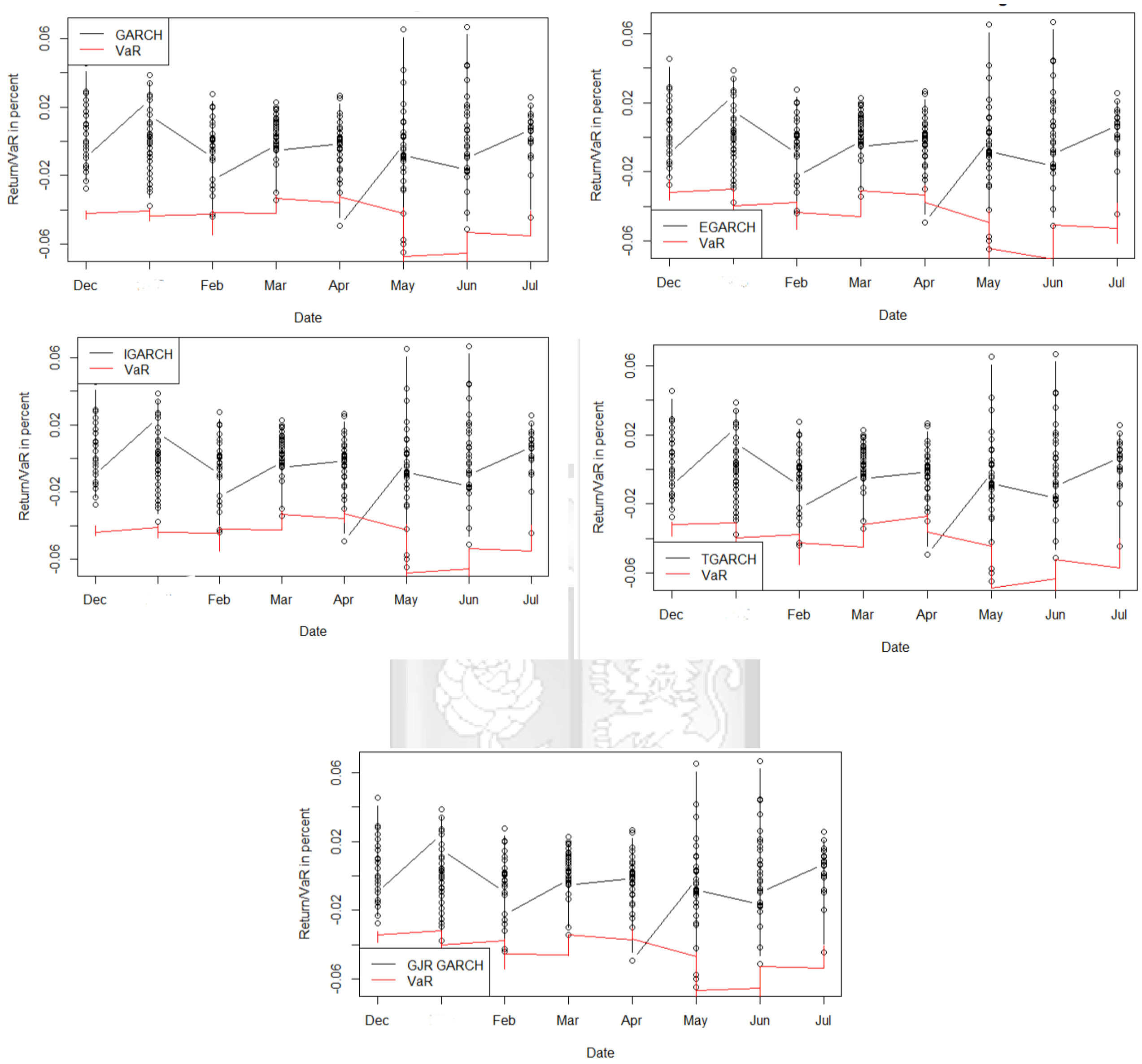

Figure 12: VaR Backtesting

The plots show that the VaR results are lower than returns from the five GARCH models. This is a clear indication that the models were fitted correctly.

4.7 Forecasting

Since we have now confirmed that our risk models works reasonably well, we proceed to produce returns forecasts. Our data has 2530 log returns. Out of these, 2024 were used for model fitting. We now use the remaining 506 for comparison with forecasted values. We will therefore forecast 506 values using the five GARCH models. To compare the accuracy of the models we get the difference between the realized values and forecasted values to get the errors. The results are as shown in Table 14;

Table 14: Errors summary - Normal Distribution

.	GARCH	EGARCH	IGARCH	TGARCH	GJR GARCH
ME	0.0002444267	0.0007117248	0.0002477164	0.0006844439	0.0005410619
RMSE	0.05350485	0.05335824	0.05350445	0.05337032	0.053378
MAE	0.0256712	0.02540721	0.02567104	0.02542058	0.02539998
MPE	NaN	NaN	NaN	NaN	NaN
MAPE	Inf	Inf	Inf	Inf	Inf

Table 15: Errors summary - Student T Distribution

.	GARCH	EGARCH	IGARCH	TGARCH	GJR GARCH
ME	0.00009929977522	0.0003638584	0.00009635908396	0.0003084351	0.0002731613
RMSE	0.05304368	0.0532057	0.05304335	0.05319495	0.0533035
MAE	0.02505944	0.02527749	0.0250589	0.02526921	0.0252244
MPE	NaN	NaN	NaN	NaN	NaN
MAPE	Inf	Inf	Inf	Inf	Inf

From the above tables, all the five models produce relatively good results. This shows that we cannot discredit the importance of any of the models in forecasting crude oil prices. However, IGARCH model outperforms all models with the simple GARCH model following closely. This shows that even with the development of new GARCH extension models, simple GARCH model is still relevant in modelling crude oil prices. The IGARCH - T distribution model has less errors than IGARCH - normal distribution. We threfore conclude that the IGARCH- T distribution is the best for Brent Crude Oil spot prices forecasting.

5 Discussion, Conclusion and Recommendations

Discussion

We set out to analyse different volatility models and their ability to forecast crude oil price volatility. The study considered five volatility models due to their ability to capture various properties of data series. The models analysed in this study are; GARCH, EGARCH, IGARCH, TGARCH AND GJR GARCH models. Brent Crude Oil spot prices data was used in our study. First, analysis of the logreturns was done. We found that the returns almost fit a normally distribution but have heavy tails. This led us to fitting the T- distribution in addition to the normal distribution. The T- distribution outperformed the normal distribution. This is a clear indication that T- distribution is better than normal distribution when the data in question has heavy tails/ outliers. Presence of mean reversion and heteroscedasticity was also verified through plots and ARCH Langrangian Multplier test. These properties indicated that GARCH models were the best to be used for our study.

Fitting of the GARCH models was done to get the parameters. The higher the logarithm maximum likelihood value, the better the fitting of the model and the lower the information criteria, the better the model. The EGARCH model is the best fitted model based on the loglikelihood value of 5446.416 and AIC value of -5.3700 out of the fitted models. We then proceeded to do backtesting VaR test which indicated that all the five models were fitted correctly. However, of the five models, IGARCH- T distibution model was the best for VaR estimations based on the LR.uc Statistic of 0.235 and LR.cc Statistic of 0.317. We therefore concluded that the IGARCH-T distribution is the best for VaR estimations for Brent Crude Oil prices.

After verifying that all the models work correctly, we proceeded and forecasted Brent crude oil daily prices. The forecasted log-returns were compared with the realized log-returns for the same period. All the GARCH model give relatively the same results for the error reports. However, the IGARCH model outperformed the other GARCH models across. The IGARCH- T distribution gave the best estimates in all the tests done based on the low ME value of 0.0000963591 and RMSE of 0.05304335. We therefore conclude that the IGARCH-T distribution model is the best for forecasting Brent Crude Oil Spot prices.

Conclusion

From our study, we find that the fact that the EGARCH model is the best fitted model based on the highest log-likelihood value of 5446.416 and least AIC value of -5.3700 in this study does not translate to the model being the best for VaR estimation and volatility forecasting. The IGARCH model is the best for both VaR estimation based on the lowest LR.uc Statistic of 0.235 and lowest LR.cc Statistic of 0.317.and volatility forecasting based on the low ME value of 0.0000963591 and RMSE of 0.05304335 for Brent Crude Oil prices. This shows the importance of testing the models performance in all stages i.e. best fitted model, backtesting, forecasting performance so as to pick the best model.

From the above findings, we can conclude that; The VaR is still a useful tool in risk management. Whenever one wants to forecast the value at risk for a commodity of a company in a long horizon, it is always better to compare all of models to choose an appropriate one, as there is hardly any model that fits a commodity forever.

Even for the same commodity of same country/companies, in different time periods, an appropriate model to predict its future VaR may vary too. Therefore, companies should choose a better risk management model based on the statistical properties of the time series, on certain time period.

The distribution used in modelling matters. From our study, we used the normal distribution and T- distribution. The T-distribution performed better across all the five models fitted. This was mainly attributed to the presence of heavy tail in the data. Therefore, the choice of the distribution is important in ensuring that we reduce forecasting errors.

Recommendation

Further research is necessary and it would be a good idea to expand the type of volatility models used. Models such as multifractal models can be considered.The number of distributions used can also be increased to include distributions such as skewed Student-t and reparameterised Johnson distributions. We also recommend tests to be done with 5% level of significance.

The recommendations for practitioners is that investors should use both VaR estimation and forecasting performance of volatility models when evaluating investment risk. VaR almost always overestimates the risk. From the forecasting results, all models give almost similar errors indicating that all the models can be relied upon to give a close estimate of future prices. However, the study shows that IGARCH model gives better results for VaR estimations and volatility forecasting.

6 Bibliography

Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in Petrov, B. N.; Csáki, F. (eds.), 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, pp. 267–281. Republished in Kotz, S.; Johnson, N. L., eds. (1992), Breakthroughs in Statistics, I, Springer-Verlag, pp. 610–624.

Andrea Bastianin & Matteo Manera (2015), How does stock market volatility react to oil price shocks?

Angelidis, T., Benos, A., & Degiannakis, S. (2004). The use of GARCH models in VaR estimation.Statistical Methodology, 1, 105–128. doi:10.1016/j.stamet.2004.08.004.

Ayhan Demirbas, Basil Omar Al-Sasi & Abdul-Sattar Nizami (2017) Recent volatility in the price of crude oil; volume 12,2017-Issue 5, Pages 408-414

Backus, D. K. and M. J. Crucini (2000). Oil prices and the terms of trade. Journal of International Economics 50, 185–213.

Blanchard, O.J. and J. Galí (2009) 'The macroeconomic effects of oil price shocks: why are the 2000s so different from the 1970s?'. In: Galí, J. and M.J. Gertler (eds.), International Dimensions of Monetary Policy, Chicago: University of Chicago Press, 373-421.

Bollerslev, T. Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327, (1986).

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (2007) TimeSeries Analysis: Forecasting and Control. 3rd edition. Pearson Education, Delhi.

Bucevska, V. (2012). An Empirical evaluation of GARCH models in value-at-risk estimation: Evidence from the Macedonian stock exchange. Business Systems Research, 4, 49–64. doi:10.2478/bsrj-2013- 0005

Chaker Aloui, Samir Mabrouk (2010) Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed GARCH models, Energy Policy 38 (2010) 2326-2339

Chortareas, G., and E., Noikokyris (2014) 'Oil shocks, stock market prices, and the U.S. dividend yield', International Review of Economics and Finance 29, 639-49.

David A. Dickey & Wayne A. Fuller (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association, 74:366a, 427-431, DOI: 10.1080/01621459.1979.10482531

Degiannakis, S., G. Filis, and R. Kizys (2014) 'The effects of oil price shocks on stock market volatility: evidence from European data', The Energy Journal, 35, 35-56.

Elder, J. and A. Serletis (2010). Oil price uncertainty. Journal of Money, Credit and Banking 42, 1137–1159.

Engle, R."Autorregressive Conditional Heteroskedasticity with Estimates of United Kingdom Inflation", Econometrica, 50, 987-1008 (1982).

Fuller, W. A. (1976). Introduction to Statistical Time Series. New York: John Wiley and Sons. ISBN 0-471-28715-6.Ogunc, A. & Hill, C. (2008) Using Excel: Companion to Principles of Econometrics, Third Edition.

James S. Doran & Ehud I. Ronn (2007) Computing the Market Price of Volatility Risk in the Energy Commodity Markets

Jarque, Carlos M.; Bera, Anil K. (1980). "Efficient tests for normality, homoscedasticity and serial independence of regression residuals". Economics Letters. 6 (3): 255–259.doi:10.1016/0165-1765(80)90024-5.

Jarque, Carlos M.; Bera, Anil K. (1981). "Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence". Economics Letters. 7 (4): 313–318.doi:10.1016/0165-1765(81)90035-5.

Jarque, Carlos M.; Bera, Anil K. (1987). "A test for normality of observations and regression residuals". International Statistical Review. 55 (2): 163–172. JSTOR 1403192

Jeff Fleming & Barbara Osrdiek The Impact of Energy Derivatives on the Crude Oil Market

Jui-Cheng Hung, Ming-Chih Lee, Hung-Chun Liu (2008) Estimation of value-at-risk for energy commodities via fat-tailed GARCH models, Energy Economics 30 (2008) 1173-1191

Julija Cerović Smolović, Milena Lipovina-Božović & Saša Vujošević (2017) GARCH models in value at risk estimation: empirical evidence from the Montenegrin stock exchange, Economic Research-Ekonomska Istraživanja, 30:1, 477-498, DOI: 10.1080/1331677X.2017.1305773

Jung, H. and C. Park (2011) 'Stock market reaction to oil price shocks: a comparison between an oil-exporting economy and an oil-importing economy,' Journal of Economic Theory and Econometrics 22, 1-29.

Kenneth Gitonga Kimathi (2018), Valuation of a locational spread option: the case of tomatoes in Nairobi and Mombasa Counties in Kenya

Khindanova, I. Atakhanova, Z. Rachev, S. GARCH-Type Processes in Modeling Energy Prices. Handbook of Computational and Numerical Methods in Finance, pp 71-110 (2004).

Klein, Judy L. 1997. Statistical Visions in Time: A History of Time Series Analysis, 1662-1938. Cambridge, U.K.: Cambridge University Press.

Leo Smigel (2021) Analysing Alpha - What is Stationarity in Time Series Analysis? A visual guide.

Lux, Thomas; Segnon, Mawuli; Gupta, Rangan (2015) : Modeling and forecasting crude oil price volatility: Evidence from historical and recent data, FinMaP-Working Paper, No. 31, Kiel University, FinMaP - Financial Distortions and Macroeconomic Performance, Kiel

Miletic, M. & Miletic, S. (2015). Performance of value at risk models in the midst of the global financial crisis in selected CEE emerging capital markets. Economic Research-Ekonomska Istraživanja, 28, 132–166. doi:10.1080/1331677X.2015.1028243.

Ming Lei Liu, Qiang Ji, Ying,Fan (2013) How does oil market uncertainty interact with other markets? An empirical analysis of implied volatility index; Volume 55, 15 June 2013, Pages 860-868

Nelson, D.B. (1991) Conditional heteroscedasticity in asset returns: A new approach. Econometrica, 59: 347-370.

Orhan, M. & Köksal, B. 2012; 2011, "A comparison of GARCH models for VaR estimation", Expert Systems with Applications, vol. 39, no. 3, pp. 3582- 3592.

Qiang Ji, Ying Fan (2012) How does oil price volatility affect non-energy commodity markets?; Volume 89, Issue 1, January 2012, Pages 273-280

Said S.E. and Dickey, D.A. (1984): "Testing for unit roots in autoregressive moving average models of unknown order." Biometrica 71,599-607

Schwert, G. W. (1989). Why does stock market volatility change over time? Journal of Finance, 44, 1115-1154.

Sergiy Ladokhin (2009), Forecasting Volatility in the stock Market

Vlaar, P.J.G. 2000, "Value at risk models for Dutch bond portfolios", Journal of banking & finance, vol. 24, no. 7, pp. 1131-1154.

Wei, Y., Wang, Y., & Huang, D. (2010). Forecasting crude oil market volatility: Further evidence using GARCH-class models. Energy Economics, 32, 1447–1484.

Wei, Y., Y. Wang, and D. Huang (2010). Forecasting crude oil market volatility: further evidence using GARCH-class models. Energy Economics 32, 1477–1484.

Ying Fan, Yue-Jun Zhang, Hsien-Tang Tsai, Yi-Ming Wei (2008) Estimating Value at Risk of crude oil price and its spillover effect using the GED-GARCH approach, Energy Economics 30 (2008) 3156-3171.

Yudong Wang a, Chongfeng Wub & Li Yang (2016), Forecasting crude oil market volatility: A Markov switching multifractal volatility approach, International Journal of Forecasting 32 (2016) 1-9

Zakoian, J. M. (1994), “Threshold heteroskedastic models,” Journal of Economic Dynamics and Control, 18(5), 931-955.

7 Appendix

7.0.1 Code

```
   ```{r}
#getting a feel for the data
brent<-read.csv("C:/Users/admin/Desktop/brentdata.csv",header=T)
#brent<-read.csv("brentdata.csv",header=T)#importing the data
#importing the data
#getting a feel for the data
str(brent)
str(brent)
head(brent)
tail(brent)
class(brent$date)
class(brent$price)
brentprices<-brent$price
plot(brent$price, type = "l", col = "blue", lwd = 2, xlab="date" ,
     ylab = "Price", main = "Daily Brent Crude Oil closing price",sub="fig1.brentdailyprices")
legend(x = 'topright', legend = 'Brent Daily prices', lty = 1, lwd = 2, col = 'blue')
qqnorm(brent$price,main="BRENT SPOT PRICES",sub="fig2.brentdailyprices")
qqline(brent$price)
```


```{r}
#to obtain returns
logreturns<-diff(log(brentprices))
#write.csv(logreturns,"C:\\Users\\admin\\Desktop\\logreturns.csv", row.names = FALSE)
meanreturn<-mean(logreturns)
meanreturn
varreturn<-var(logreturns)
varreturn
sdreturn<-sd(logreturns)
sdreturn
library(moments)
skew<-skewness(logreturns)
skew
kurt<-kurtosis(logreturns)
kurt
```

#we look at the histogram and/or the empirical distribution of daily Brent log returns and compare it with the normal distribu
#We use the function curve()with an additional parameter add=TRUE to plot a second line to an already existing diagram

```{r}
#lreturns<-read.csv("C:/Users/admin/Desktop/lreturns.csv",header=T)
par(mfrow=c(1,2))
hist(logreturns, nclass=40, freq=FALSE, main='Log Return histogram',sub="fig5. Comparison with normal distribution");curve(dnorm
plot(density(logreturns), main='Log Return empirical distribution');curve(dnorm(x, mean=meanreturn,sd=sdreturn), from = -0.3, to
par(mfrow=c(1,1))

plot(logreturns, type = "l", col = "blue", lwd = 2, xlab="date" ,
     ylab = "log returns", main = "Daily Brent Crude Oil returns", sub="fig3.brentreturns")
legend(x = 'topleft', legend = 'Brent Log Returns', lty = 1, lwd = 2, col = 'blue')

squaredreturns<-logreturns^2
plot(squaredreturns, type = "l", col = "blue", lwd = 2, xlab="date" ,
     ylab = "squared log returns", main = "Daily Brent Crude Oil squared log returns" , sub="fig4.brentsquaredreturns")
legend(x = 'topleft', legend = 'Brent Squared Log Returns', lty = 1, lwd = 2, col = 'blue')
```
#ARCH Lagrangian Multiplier test
```{r}
library(nortsTest)
arch.test(logreturns,arch="box",alpha=0.05,lag.max = 2)
```

density plots on log-scale
```{r}
plot(density(logreturns), xlim=c(-5*sdreturn,5*sdreturn),log='y', main='Density on log-scale')
curve(dnorm(x, mean=meanreturn,sd=sdreturn), from=-5*sdreturn, to=5*sdreturn, log="y", add=TRUE, col="red")
```
```

````
# QQ-plot
```{r}
qqnorm(logreturns,main="BRENT RETURNS",sub="fig6.returns")
qqline(logreturns)
```


##Jarque Bera Test
```{r}
jarque <- arima(logreturns, order = c(0, 1, 0), seasonal = list(order = c(2, 0, 1)))
jarque.bera.test(residuals(jarque))
jarque.bera.test(logreturns)

```


```{r}
#Stationarity test using ADF
library(tseries)
library(urca)
adf_brent = summary(ur.df(diff(logreturns), type = "trend", selectlags = "BIC"))
adf_brent
ur.df(logreturns, type = c("none", "drift", "trend"), lags = 1, selectlags = c("Fixed", "AIC", "BIC"))
adf.test(logreturns, alternative = c("stationary", "explosive"),
         k = trunc((length(logreturns)-1)^(1/3)))
```

```{r}
#DETERMINING MEAN EQUATION
model1<-arima(logreturns[1:2024],order=c(4,0,3))
model1
model2<-arima(logreturns[1:2024],order=c(4,0,4))
model2
model3<-arima(logreturns[1:2024],order=c(4,0,2))
model3
model4<-arima(logreturns[1:2024],order=c(3,0,3))
model4
model5<-arima(logreturns[1:2024],order=c(5,0,3))
model5
model6<-arima(logreturns[1:2024],order=c(0,0,5))
model6
model7<-arima(logreturns[1:2024],order=c(5,0,0))
model7
model8<-arima(logreturns[1:2024],order=c(0,0,4))
model8
model9<-arima(logreturns[1:2024],order=c(4,1,3))
model9
model10<-arima(logreturns[1:2024],order=c(4,1,4))
model10
model11<-arima(logreturns[1:2024],order=c(4,1,2))
model11
model12<-arima(logreturns[1:2024],order=c(2,1,3))
model12
model13<-arima(logreturns[1:2024],order=c(5,1,3))
model13
model14<-arima(logreturns[1:2024],order=c(4,0,0))
model14
model15<-arima(logreturns[1:2024],order=c(3,0,4))
model15
model16<-arima(logreturns[1:2024],order=c(2,0,4))
model16
model16<-arima(logreturns[1:2024],order=c(2,0,3))
model16
model17<-arima(logreturns[1:2024],order=c(0,0,0))
model17
model18<-arima(logreturns[1:2024],order=c(5,0,5))
model18
```

##supporting the findings
```{r}
library(forecast)
auto.arima(logreturns[1:2024])
```
````

```
#Testing for ARCH effects
```{r}
par(mfrow=c(2,1))
acf(logreturns[1:2024])
pacf(squaredreturns[1:2024],sub="fig7.ACF,PACF")
pacf(logreturns[1:2024],sub="fig7.ACF,PACF")
residuals<-(model1$residuals)
residuals2<-residuals^2
acf(residuals)
pacf(residuals2,sub="fig8.residuals ACF,PACF")
F<-abs(residuals)
library(FinTS)
ArchTest(logreturns[1:2024],lag=10,demean=FALSE)
```


#Fitting GARCH Models
```{r}
library(forecast)
library(fGarch)
library(rugarch)
library(quantmod)
library(xts)
library(forecast)
library(rugarch)
library(fGarch)
library(tseries)
library(ggplot2)
#df2<-df[2:2531]
T <- nrow(lreturns)-1
T
T_train <- round(4/5*T)
T_train
T_test <- T - T_train
T_test
dates_out_of_sample <- tail(index(logreturns), T_test)
dates_all <- index(logreturns)
dates_in_sample <- dates_all[1:T_train]
```
#simple GARCH (1,1) model, (sGARCH)
```{r}
garch11spec<-ugarchspec(variance.model=list(model = "sGARCH",garchOrder=c(1, 1),
                                            submodel=NULL,external.regressors=NULL,variance.targeting=FALSE),
                        mean.model=list(armaOrder = c(2,4),include.mean=TRUE,archm=FALSE,
                                        archpow=1,arfima=FALSE,external.regressors=NULL,archex=FALSE),
                        distribution.model="norm")
garch11spec
```

#fitting this model to our data -estimate the unknown parameters by MLE, based on our time series of daily returns
```{r}
garch11fit<-ugarchfit(data=logreturns,spec = garch11spec, out.sample = T_test)
show(garch11fit)
coef(garch11fit)
```

#creating a news impact curve
```{r}
newsimpactgarch11 <- newsimpact(garch11fit)
plot(newsimpactgarch11$zx, newsimpactgarch11$zy, type="l", lwd=2, col="blue", main="GARCH(1,1) - News Impact", ylab=newsimpactga
```


#EGARCH model
specify EGARCH(1,1) model
```{r}
egarch11spec <- ugarchspec(variance.model=list(model = "eGARCH",garchOrder=c(1,1),
                                               submodel=NULL,external.regressors=NULL,variance.targeting=FALSE),
                           mean.model=list(armaOrder = c(2,0,4),include.mean=TRUE,archm=FALSE,
                                           archpow=1,arfima=FALSE,external.regressors=NULL,archex=FALSE),
                           distribution.model="norm")
egarch11spec
```

#fitting the EGARCH model
```{r}
egarch11fit <- ugarchfit(data=logreturns,spec = egarch11spec,out.sample = T_test)
```
```

```
show(egarch11fit)
coef(egarch11fit)
```

#creating a news impact curve
```{r}
newsimpactegarch11 <- newsimpact(egarch11fit)
plot(newsimpactegarch11$zx, newsimpactegarch11$zy, type="l", lwd=2, col="blue", main="EGARCH(1,1) - News Impact", ylab=newsimpa
```

##IGARCH MODEL
specify IGARCH(1,1) model
```{r}
igarch11spec <- ugarchspec(variance.model=list(model = "iGARCH",garchOrder=c(1, 1),
                                               submodel=NULL,external.regressors=NULL,variance.targeting=FALSE),
                           mean.model=list(armaOrder = c(2,4),include.mean=TRUE,archm=FALSE,
                                           archpow=1,arfima=FALSE,external.regressors=NULL,archex=FALSE),
                           distribution.model="norm")
igarch11spec
```

#fitting the IGARCH model
```{r}
igarch11fit <- ugarchfit(data=logreturns,spec = igarch11spec, out.sample = T_test)
show(igarch11fit)
coef(igarch11fit)
```

#creating a news impact curve
```{r}
newsimpactigarch11 <- newsimpact(igarch11fit)
plot(newsimpactigarch11$zx, newsimpactigarch11$zy, type="l", lwd=2, col="blue", main="IGARCH(1,1) - News Impact", ylab=newsimpa
```
##Threshold GARCH
specify TGARCH(1,1) model
```{r}

tgarch11spec <- ugarchspec(variance.model=list(model="fGARCH", submodel="TGARCH", garchOrder=c(1, 1),
                                               submodel=NULL,external.regressors=NULL,variance.targeting=FALSE),
                           mean.model=list(armaOrder = c(2,4),include.mean=TRUE,archm=FALSE,
                                           archpow=1,arfima=FALSE,external.regressors=NULL,archex=FALSE),
                           distribution.model="norm")
tgarch11spec
```
#fitting the TGARCH model
```{r}
tgarch11fit <- ugarchfit(data=logreturns,spec = tgarch11spec, out.sample = T_test)
show(tgarch11fit)
coef(tgarch11fit)
```
#creating a news impact curve
```{r}
newsimpacttgarch11 <- newsimpact(tgarch11fit)
plot(newsimpacttgarch11$zx, newsimpacttgarch11$zy, type="l", lwd=2, col="blue", main="TGARCH(1,1) - News Impact", ylab=newsimpa
```
###GJR GARCH model
specify GJR GARCH(1,1) model
```{r}
gjrgarch11spec <- ugarchspec(variance.model=list(model = "gjrGARCH",garchOrder=c(1, 1),
                                               submodel=NULL,external.regressors=NULL,variance.targeting=FALSE),
                           mean.model=list(armaOrder = c(2,4),include.mean=TRUE,archm=FALSE,
                                           archpow=1,arfima=FALSE,external.regressors=NULL,archex=FALSE),
                           distribution.model="norm")
gjrgarch11spec
```
#fitting the GJRGARCH model
```{r}
gjrgarch11fit <- ugarchfit(data=logreturns,spec = gjrgarch11spec, out.sample = T_test)
show(gjrgarch11fit)
coef(gjrgarch11fit)
```
#creating a news impact curve
```{r}
newsimpactgjrgarch11 <- newsimpact(gjrgarch11fit)
plot(newsimpactgjrgarch11$zx, newsimpactgjrgarch11$zy, type="l", lwd=2, col="blue", main="GJRGARCH(1,1) - News Impact", ylab=new
```

```
##Forecasting using garch models
##GARCH(1,1) model forecasting
‘‘‘{r}
garch11.specc = ugarchspec(variance.model = list(model = "sGARCH",garchOrder=c(1,1)),
                           mean.model = list(armaOrder=c(2,4)))
garch11.specc
garch11forecastfit<-ugarchfit(spec=garch11.specc, data=logreturns, out.sample=T_test)
garch11forecastfit
coef(garch11forecastfit)
‘‘‘
###carrying out historical backtesting
‘‘‘{r}
garch11back <- ugarchroll(garch11.specc, data= logreturns[1:2024], n.start = 1800, refit.every = 1, refit.window = "moving",sol
report(garch11back, type = "VaR", VaR.alpha = 0.01,conf.level = 0.99)
‘‘‘

## plot of the backtesting performance
‘‘‘{r}
backtestVaR <- zoo(garch11back@forecast$VaR[, 1])
index(backtestVaR) <- as.yearmon(rownames(garch11back@forecast$VaR))
backactual <- zoo(garch11back@forecast$VaR[, 2])
index(backactual) <- as.yearmon(rownames(garch11back@forecast$VaR))
plot(backactual, type = "b", main = "99% 1 Month VaR Backtesting", xlab = "Date", ylab = "Return/VaR in percent")
lines(backtestVaR, col = "red")
legend("topleft",c("GARCH","VaR"), col = c("black","red"), lty = c(1,1))
‘‘‘

# forecast log-returns along the whole out-of-sample
‘‘‘{r}
library(xts)
garch_fore <- ugarchforecast(garch11forecastfit,n.ahead = 505, n.roll = T_test-1)
forecast_log_returns <- xts(garch_fore@forecast$seriesFor[1, ],order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
forecast_volatility <- xts(garch_fore@forecast$sigmaFor[1, ], order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
forecast_volatility
## Getting the data
fit1<-fitted(garch11forecastfit)
fit1[1:675]
fit1[676:1350]
fit1[1351:2024]
forecast_log_returns

‘‘‘

# plot of log-returns
‘‘‘{r}
initialreturns<-read.csv("C:/Users/admin/Desktop/initialreturns.csv",header=T)
ireturns<-data.frame(initialreturns)
ireturns
garch11fitted<-read.csv("C:/Users/admin/Desktop/garch11fitted.csv",header=T)
garch11forecast<-read.csv("C:/Users/admin/Desktop/garch11forecast.csv",header=T)
plot(cbind("fitted"   = garch11fitted$garch11fitted,
           "forecast" = garch11forecast$garch11forecast,
           "original" = initialreturns$logreturns),
     col = c("blue", "red", "black"), lwd = c(0.5, 0.5, 2),type='l',
     main = "Forecast of log-returns", legend.loc = "topleft")
plot(garch11fitted$garch11fitted, type='l', col='blue')
line(garch11forecast$garch11forecast)
ggplot(ggplot(ireturns, aes(x = date, y = logreturns)))
plot(forecast_log_returns, which='all')
line(logreturns[2025:2530], col="red")
plot(forecast_log_returns)
line(logreturns)
ggplot(ireturns)
plot(logreturns, type='l')
line(garch11fitted, color="red")
‘‘‘


##EGARCH(1,1) model forecasting
‘‘‘{r}
Egarch11.specc = ugarchspec(variance.model = list(model = "eGARCH",garchOrder=c(1,1)),
                           mean.model = list(armaOrder=c(2,4)))
Egarch11forecastfit<-ugarchfit(spec=Egarch11.specc, data=logreturns, out.sample=T_test)
Egarch11forecastfit
coef(Egarch11forecastfit)
‘‘‘
```

~~~~
###carrying out historical backtesting
```{r}
Egarch11back <- ugarchroll(Egarch11.specc, data= logreturns[1:2024], n.start = 1800, refit.every = 1, refit.window = "moving",s
report(Egarch11back, type = "VaR", VaR.alpha = 0.01,conf.level = 0.99)
```

## plot of the backtesting performance
```{r}
EbacktestVaR <- zoo(Egarch11back@forecast$VaR[, 1])
index(EbacktestVaR) <- as.yearmon(rownames(Egarch11back@forecast$VaR))
Ebackactual <- zoo(Egarch11back@forecast$VaR[, 2])
index(Ebackactual) <- as.yearmon(rownames(Egarch11back@forecast$VaR))
plot(Ebackactual, type = "b", main = "99% 1 Month VaR Backtesting", xlab = "Date", ylab = "Return/VaR in percent")
lines(EbacktestVaR, col = "red")
legend("bottomleft",c("EGARCH","VaR"), col = c("black","red"), lty = c(1,1))
```

# forecast log-returns along the whole out-of-sample
```{r}
Egarch_fore <- ugarchforecast(Egarch11forecastfit, n.ahead = 505, n.roll = T_test-1)
Egarch_fore
Eforecast_log_returns <- xts(Egarch_fore@forecast$seriesFor[1, ],order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
Eforecast_log_returns
Eforecast_volatility <- xts(Egarch_fore@forecast$sigmaFor[1, ], order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
Eforecast_volatility
fit2<-fitted(Egarch11forecastfit)
fit2[1:675]
fit2[676:1350]
fit2[1351:2024]
```

##plot of log-returns
```{r}
plot(cbind("fitted"   = fitted(Egarch11forecastfit),
           "forecast" = Eforecast_log_returns,
           "original" = logreturns),
     col = c("blue", "red", "black"), lwd = c(0.5, 0.5, 2),
     main = " EGARCH Forecast of log-returns", legend.loc = "topleft")
```


##IGARCH(1,1) model forecasting
```{r}
Igarch11.specc = ugarchspec(variance.model = list(model = "iGARCH",garchOrder=c(1,1)),
                            mean.model = list(armaOrder=c(2,4)))
Igarch11forecastfit<-ugarchfit(spec=Igarch11.specc, data=logreturns, out.sample=T_test)
```

###carrying out historical backtesting
```{r}
Igarch11back <- ugarchroll(Igarch11.specc, data= logreturns[1:2024], n.start = 1800, refit.every = 1, refit.window = "moving",s
report(Igarch11back, type = "VaR", VaR.alpha = 0.01,conf.level = 0.99)
```

## plot of the backtesting performance
```{r}
IbacktestVaR <- zoo(Igarch11back@forecast$VaR[, 1])
index(IbacktestVaR) <- as.yearmon(rownames(Igarch11back@forecast$VaR))
Ibackactual <- zoo(Igarch11back@forecast$VaR[, 2])
index(Ibackactual) <- as.yearmon(rownames(Igarch11back@forecast$VaR))
plot(Ibackactual, type = "b", main = "99% 1 Month VaR Backtesting", xlab = "Date", ylab = "Return/VaR in percent")
lines(IbacktestVaR, col = "red")
legend("topleft",c("IGARCH","VaR"), col = c("black","red"), lty = c(1,1))
```

# forecast log-returns along the whole out-of-sample
```{r}
Igarch_fore <- ugarchforecast(Igarch11forecastfit, n.ahead = 505, n.roll = T_test-1)
Igarch_fore
Iforecast_log_returns <- xts(Igarch_fore@forecast$seriesFor[1, ],order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
Iforecast_log_returns
Iforecast_volatility <- xts(Igarch_fore@forecast$sigmaFor[1, ], order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
Iforecast_volatility
fit3<-fitted(Igarch11forecastfit)
fit3[1:675]
fit3[676:1350]
~~~~

~~~
fit3[1351:2024]
```

# plot of log-returns
```{r}
plot(cbind("fitted"   = fitted(Igarch11forecastfit),
           "forecast" = Iforecast_log_returns,
           "original" = logreturns),
     col = c("blue", "red", "black"), lwd = c(0.5, 0.5, 2),
     main = " IGARCH Forecast of log-returns", legend.loc = "topleft")
```
##TGARCH(1,1) model forecasting
```{r}
Tgarch11.specc = ugarchspec(variance.model = list(model="fGARCH", submodel="TGARCH",garchOrder=c(1,1)),
                            mean.model = list(armaOrder=c(2,4)))
Tgarch11forecastfit<-ugarchfit(spec=Tgarch11.specc, data=logreturns, out.sample=T_test)
```

###carrying out historical backtesting
```{r}
Tgarch11back <- ugarchroll(Tgarch11.specc, data= logreturns[1:2024], n.start = 1800, refit.every = 1, refit.window = "moving",sc
report(Tgarch11back, type = "VaR", VaR.alpha = 0.01,conf.level = 0.99)
```
```{r}
## plot of the backtesting performance
TbacktestVaR <- zoo(Tgarch11back@forecast$VaR[, 1])
index(TbacktestVaR) <- as.yearmon(rownames(Tgarch11back@forecast$VaR))
Tbackactual <- zoo(Tgarch11back@forecast$VaR[, 2])
index(Tbackactual) <- as.yearmon(rownames(Tgarch11back@forecast$VaR))
plot(Tbackactual, type = "b", main = "99% 1 Month VaR Backtesting", xlab = "Date", ylab = "Return/VaR in percent")
lines(TbacktestVaR, col = "red")
legend("bottomleft",c("TGARCH","VaR"), col = c("black","red"), lty = c(1,1))
```
# forecast log-returns along the whole out-of-sample
```{r}
Tgarch_fore <- ugarchforecast(Tgarch11forecastfit, n.ahead = 505, n.roll = T_test-1)
Tgarch_fore
Tforecast_log_returns <- xts(Tgarch_fore@forecast$seriesFor[1, ],order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
Tforecast_log_returns
Tforecast_volatility <- xts(Tgarch_fore@forecast$sigmaFor[1, ], order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
Tforecast_volatility
fit4<-fitted(Tgarch11forecastfit)
fit4[1:675]
fit4[676:1350]
fit4[1351:2024]
```

# plot of log-returns
```{r}
plot(cbind("fitted"   = fitted(Igarch11forecastfit),
           "forecast" = Iforecast_log_returns,
           "original" = logreturns),
     col = c("blue", "red", "black"), lwd = c(0.5, 0.5, 2),
     main = " IGARCH Forecast of log-returns", legend.loc = "topleft")
```
##GJR GARCH(1,1) model forecasting
```{r}
GJRgarch11.specc = ugarchspec(variance.model = list(model="gjrGARCH",garchOrder=c(1,1)),
                            mean.model = list(armaOrder=c(2,4)))
GJRgarch11forecastfit<-ugarchfit(spec=GJRgarch11.specc, data=logreturns, out.sample=T_test)
```
###carrying out historical backtesting
```{r}
GJRgarch11back <- ugarchroll(GJRgarch11.specc, data= logreturns[1:2024], n.start = 1800, refit.every = 1, refit.window = "moving
report(GJRgarch11back, type = "VaR", VaR.alpha = 0.01,conf.level = 0.99)
```
## plot of the backtesting performance
```{r}
GJRbacktestVaR <- zoo(GJRgarch11back@forecast$VaR[, 1])
index(GJRbacktestVaR) <- as.yearmon(rownames(GJRgarch11back@forecast$VaR))
GJRbackactual <- zoo(GJRgarch11back@forecast$VaR[, 2])
index(GJRbackactual) <- as.yearmon(rownames(GJRgarch11back@forecast$VaR))
plot(GJRbackactual, type = "b", main = "99% 1 Month VaR Backtesting", xlab = "Date", ylab = "Return/VaR in percent")
lines(GJRbacktestVaR, col = "red")
legend("bottomleft",c("GJR GARCH","VaR"), col = c("black","red"), lty = c(1,1))
```
# forecast log-returns along the whole out-of-sample
~~~

```
‘‘‘{r}
GJRgarch_fore <- ugarchforecast(GJRgarch11forecastfit, n.ahead = 505, n.roll = T_test-1)
GJRgarch_fore
GJRforecast_log_returns <- xts(GJRgarch_fore@forecast$seriesFor[1, ],order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
GJRforecast_log_returns
GJRforecast_volatility <- xts(GJRgarch_fore@forecast$sigmaFor[1, ], order.by = as.Date(brent$date[2026:2531], "%m/%d/%Y"))
GJRforecast_volatility
fit5<-fitted(GJRgarch11forecastfit)
fit5[1:675]
fit5[676:1350]
fit5[1351:2024]
‘‘‘
# plot of log-returns
‘‘‘{r}
plot(cbind("fitted"   = fitted(Igarch11forecastfit),
           "forecast" = Iforecast_log_returns,
           "original" = logreturns),
     col = c("blue", "red", "black"), lwd = c(0.5, 0.5, 2),
     main = " IGARCH Forecast of log-returns", legend.loc = "topleft")
‘‘‘
accuracy(GJRforecast_log_returns, logreturns[2025:2530])
‘‘‘{r}
realizedreturns<-read.csv("C:/Users/admin/Desktop/realized returns.csv",header=T)
garch11forecast1<-read.csv("C:/Users/admin/Desktop/garch11forecast.csv",header=T)
egarch11foorecast1<-read.csv("C:/Users/admin/Desktop/egarch11forecast.csv",header=T)
igarch11forecast1<-read.csv("C:/Users/admin/Desktop/igarch11forecast.csv",header=T)
tgarch11forecast1<-read.csv("C:/Users/admin/Desktop/tgarch11forecast.csv",header=T)
gjrgarch11forecast1<-read.csv("C:/Users/admin/Desktop/gjrgarch11forecast.csv",header=T)
g11forecast<-ts(garch11forecast1)
realized<-ts(realizedreturns)
eg11forecast<-ts(egarch11foorecast1)
ig11forecast<-ts(igarch11forecast1)
tg11forecast<-ts(tgarch11forecast1)
gjr11forecast<-ts(gjrgarch11forecast1)
garch11accuracy<-accuracy(g11forecast,realized)
garch11accuracy
egarch11accuracy<-accuracy(eg11forecast,realized)
egarch11accuracy
igarch11accuracy<-accuracy(ig11forecast,realized)
igarch11accuracy
tgarch11accuracy<-accuracy(tg11forecast,realized)
tgarch11accuracy
gjrgarch11accuracy<-accuracy(gjr11forecast,realized)
gjrgarch11accuracy
‘‘‘
```

7.0.2 Similarity Check

Curiginal

Document Information

Analyzed document	THESIS.pdf (D109358304)
Submitted	6/20/2021 6:40:00 PM
Submitted by	
Submitter email	Faith.Wacuka@strathmore.edu
Similarity	12%
Analysis address	library.strath@analysis.urkund.com

Sources included in the report

W	URL: https://palomar.home.ece.ust.hk/MAFS6010R_lectures/Rsession_time_series_modeling.html Fetched: 4/26/2020 8:12:15 PM	11
W	URL: https://www.researchgate.net/publication/305259468_The_one-step_ahead_Time-varying_density_forecast_window_of_fat-tailed_Value-at-risk_models Fetched: 5/4/2021 12:23:17 PM	2
W	URL: http://web.vu.lt/mif/a.buteikis/wp-content/uploads/2018/02/Lecture_02.pdf Fetched: 6/20/2021 6:41:00 PM	1
W	URL: https://www.zora.uzh.ch/id/eprint/163936/1/20100939_003325181.pdf Fetched: 6/20/2021 6:41:00 PM	1
W	URL: https://math.stackexchange.com/questions/3922710/particular-solution-of-t2y-2y-t2-using-the-method-of-variation-of-paramete Fetched: 6/20/2021 6:41:00 PM	1
W	URL: https://www.radford.edu/~thompson/webddes/tutorial.html Fetched: 6/20/2021 6:41:00 PM	1
W	URL: https://arxiv.org/pdf/1812.05985 Fetched: 6/20/2021 6:41:00 PM	1
W	URL: https://escholarship.org/content/qt7q60s3nz/qt7q60s3nz_noSplash_03a06e2a0b595f17c4709 91690d63c11.pdf Fetched: 6/20/2021 6:41:00 PM	1
SA	**Askvik & Vallenå.docx** Document Askvik & Vallenå.docx (D10823631)	11
SA	**2300252_ ACCFIN5008P_ Dissertation and Research Methods..docx** Document 2300252_ ACCFIN5008P_ Dissertation and Research Methods..docx (D40889547)	1
W	URL: https://rstudio-pubs-static.s3.amazonaws.com/164446_e77902b2248f489ca8e69d0713c8350e.html Fetched: 12/9/2019 5:18:53 PM	1
W	URL: https://www.fundacionmapfre.org/documentacion/publico/i18n/catalogo_imagenes/imagen_id.cmd?idImagen=1107112	2

1/36

Fetched: 1/17/2020 6:14:12 PM

W	URL: https://cran.r-project.org/web/packages/rugarch/rugarch.pdf Fetched: 6/20/2021 6:41:00 PM	6
W	URL: http://hedibert.org/wp-content/uploads/2018/05/garchmodeling-Rmd.txt Fetched: 1/14/2021 7:23:12 PM	1
SA	**R script - Group R 04764.pdf** Document R script - Group R 04764.pdf (D69307243)	7
SA	**RStudio code Group J.docx** Document RStudio code Group J.docx (D69307384)	1
SA	**Group X_Rcode.docx** Document Group X_Rcode.docx (D69119498)	1
W	URL: https://rstudio-pubs-static.s3.amazonaws.com/258811_b43d4c7bb2c74851b5b95f29a09c5b30.html Fetched: 12/6/2019 3:57:59 PM	1
W	URL: https://tianxie88.files.wordpress.com/2014/06/project-by-tian-xie-and-zizhen-li.pdf Fetched: 4/27/2020 11:09:42 AM	1