# ABSTRACT

The study investigated volatility clustering, leverage effects and long range dependence of British Petroleum compound returns from 1998 to 2018. The volatility clustering, leverage effects and long range dependence of returns are fitted by symmetric, asymmetric and long memory Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. The Akaike Information Criteria and Bayesian Information Criteria were used in selecting distribution of innovation for GARCH-type model among eight distributions that are able to capture heavy tailed returns. The Standard GARCH(1,1) assuming Johnson's SU distribution, Exponential GARCH (1,1) assuming Johnson's SU distribution and Fractional Integrated GARCH (1,1) assuming student-t distribution were best fit for symmetric, asymmetric and long memory GARCH-type models respectively. The Threshold GARCH (1,1) assuming Johnson's SU distribution proved to be the most stable model foe estimating Value at Risk (VaR). The best model in forecasting the stylized facts of British Petroleum returns was Fractional Integrated GARCH (1,1) assuming student-t distribution due to the minimum values of Mean Absolute, Mean Square Error and Root Mean Square Error.

# ACKNOWLEDGEMENTS

I would like to thank my Supervisor Dr F. Matarise and Prof E. Ranganai for helping me in this study and further pushing me in doing what is expected of masters students. My sincere gratitude goes to Mr L. Dhliwayo for his support especially with innovative ideas. Finally I want to thank all my classmates and staff members of the department for the encouragement concerning my dissertation.

# Chapter 1

# INTRODUCTION

Financial times series exhibit stylized facts such as volatility clustering, leverage effects, long range dependence and heavy tailed returns. Volatility is exhibited if there are any changes in conditional variance over time, this phenomena is called heteroscedasticity. Volatility clustering is recognised when large shocks tend to be followed by large shocks and small shocks tends to be followed by small shocks. In simple terms, volatility clustering exhibit itself as quite periods interrupted by volatile periods called turbulence.

Leverage effects are recognised when stock prices tends to be larger for a fall in stock prices than rise in stock price. Long range dependence is defined as the existence of dependence realised by hyperbolic decay of squared returns. Returns are not normal distributed because of the presence of heavy tailed returns more than of Gaussian distribution. Modelling stylized facts is used to quantify Value at Risk (VaR) associated with returns over a specified period.

## 1.1 Background of the Study

The stock returns exhibit changes in variance over time, hence the assumption of constant variance (homoscedasticity) is violated. In financial time series literature, models have been developed which captures the changes in conditional variance (volatility). Engle (1982) developed Autoregressive Conditional Heteroscedasticity (ARCH), which models the dynamic behaviour of conditional variance depending on squared disturbance. However, the ARCH model has its short comings that it has infinite number of parameters. Bollerslev (1986) extended the ARCH models to Generalized ARCH (GARCH) to model conditional variance, which depends simultaneous on squared disturbances and past conditional variance, which reduces the number of parameters. GARCH are classified into three namely symmetric, asymmetric and long memory GARCH models.

Symmetric GARCH models have the ability to capture volatility clustering and partly heavy tailed returns. Symmetric GARCH models are standard GARCH (sGARCH) model proposed by Bollerslev (1986), Integrated GARCH (IGARCH) model of Engle and Bollerslev (1986) and GARCH-in-Mean (GARCH-M) model of Engle, Lilien and Robins (1987). However, symmetric GARCH models cannot accommodate leverage effects and long range dependence of returns .

Asymmetric GARCH models have the ability to capture both volatility clustering and leverage effects. Asymmetric GARCH models include Exponential GARCH (EGARCH) model of Nelson (1991), Power GARCH (PGARCH) model of Ding, Engle and Granger (1993) and Threshold GARCH (TGARCH) model of Zakoin (1994). However, the short coming of asymmetric GARCH models does not embrace long range dependence of returns.

The short comings of symmetric and asymmetric GARCH models motivated econometricians to develop models that accommodate volatility clustering, leverage effects and long range dependence of conditional variance. Long memory GARCH models include Fractionally Integrated (FIGARCH) of Baille el al (1996), Fractionally Integrated Exponential GARCH (FIEGARCH) model of ......... and Adaptive Fractionally Integrated (AFIGARCH) of Baille et al and Moran (2009).

## 1.2 Aim

The aim of this study is to forecast financial stylized facts associated with stock returns using GARCH-type models assuming, namely normal distribution, skewed normal distribution, student-t distribution, skewed student-t distribution, generalized error distribution, skewed generalized error distribution, normal inverse Gaussian distribution and Johnson's SU distribution as distributions of innovations.

## 1.3 Objectives

1. To model volatility clustering of returns using symmetric GARCH-type models.
2. To model leverage effects of returns using asymmetric GARCH-type models.
3. To model long range dependence of returns using long memory GARCH-type models.
4. To estimate Value at Risk (VaR) of returns using GARCH-type models.

5. To compare the forecasting power of symmetric, asymmetric and long memory GARCH-types models in forecasting stylized facts.

## 1.4 Problem Statement

Volatility clustering has an adversely negative impact on investor's capital asset. Leverage effects increases the uncertainty to the investor and financial markets.

## 1.5 Significance of the Study

Modelling volatility clustering is essential element in portfolio management, risk management and price equity. It is key to obtain estimates of models that capture the leverage effects in order to improve portfolio allocation and valuation of financial derivatives. Capturing the impact of long range dependence in modelling is of paramount importance, in that it allows enhance stability and reassure investors and financial markets.

## 1.6 The Layout of the Study

This thesis is divided into five chapters as follows; Chapter 1: contains the Introduction, Background of the study, Aim, Objectives, Problem Statement and significance of the study. Chapter 2, contains literature on Generalised Autoregressive Conditional Heteroskedasticity(GARCH) model , Value at Risk, distributions of innovations and forecasting power methods.Chapter 3, provides methods that are used to model heavy tailed returns using GARCH-type models and distribution of innovations. Chapter 4, data analysis contains results analysis for GARCH-type models and distributions innovations. Chapter 5: Discussion of results, conclusions and recommendations.

# Chapter 2

# LITERATURE REVIEW

### 2.0.1 Introduction

The chapter discusses relevant literature concerned with modeling heavy tailed returns using GARCH-type models and distribution of innovation.

### 2.0.2 Time Series Analysis

Time series is a sequence of observation, measured typically at successive time instants spaced at uniform time interval. The study of time series is significant in identifying and characterising the nature of the data and forecasting.

#### Components of Time Series

Time series can be decomposed into the following components:

- **Trend:** A trend refers to an upward and downward movement in the data over a period of time.
- **Cyclic Component:** A cyclic component pattern refers to the recurring up and down over a period of time, which is not fixed. The fluctuations can have

a duration of at least two years and measured from peak to peak.

- **Seasonal Component:**A seasonal variation are patterns in time series that is influenced by seasonal factors (e.g the quarter of the year, the month, or day of the week). Seasonality is always of a fixed and known period.
- **Irregular Component:** Irregular fluctuations are erratic movements in a time series that does not follow any pattern. Many irregular fluctuations are caused by usual events.

### 2.0.3 Stationarity

Dasgupa (2012) defined a stationarity time series if its mean and variance are constant over a period of time. A significant number of studies have reflected that most financial time series data is non stationary with deterministic component, Gujarati and Porter (2008) suggested that estimating time series models with non stationary variables leads to biased of estimates and inefficient forecasting.
Therefore, need may arise to transform the time series data to achieve stationarity. To eradicate spurious results, enhance the accuracy of estimation of parameters and forecasting . The various technique which can be used for stationarity of the time series data are differencing, logarithmic transformation, square root transformation, reciprocal transformation and power transformation.

- **Dickey-Fuller (DF) Test**

  Dickey and Fuller (1979) developed DF test to determine whether the unit root is present in AR(1) model.

  Consider an AR(1) model $y_t = by_{t-1} + \epsilon_t$. The unit root hypothesis

$$H_0 = b = 1 \quad vs \quad H_1 < 1$$

$$Test \quad Statistics: \quad Z = \frac{\hat{b} - 1}{s.e(\hat{b})}$$

  The asymptotic distribution of Z is not normal. The hypothesis of a unit root is rejected for negative values of Z.

- **Augmented Dickey-Fuller (ADF) Test**

  Dickey and Fuller () extended DF test to a more comprehensive test called

ADF for stationarity which accomodates errors that are autocorrelated. The ADF test extents to AR(p) model.

- **Kwiatkowski-Philips-Schmidt-Shin (KPSS) Test**

  Kwiatkowski, Philips,Schmidt and Shin (1992) proposed KPSS test that is used for stationarity around a deterministic component. They derive their test statistic by starting with model:

$$y_t = \beta' D_t + \mu_t + U_t \mu_t$$

$$y_t = \mu_{t-1} + \epsilon_t \quad \epsilon_t \sim N(0, \sigma_t^2)$$

  where $D_t$ contains deterministic component, $U_t$ is $I(0)$ and may be heteroscedastic in nature and $\mu_t$ is a random walk with innovation variance $\sigma_\epsilon^2$.
  The hypothesis $H_0 : \sigma_t^2 = 0$
  which means that $\mu_t$ is constant. The KPSS test statistic is defined as

$$KPSS = \frac{T^{-2} \sum_{t=1}^{T} \hat{S}_T^2}{\hat{\lambda}^2}$$

  where $S_t = \sum_{j=1}^{t} \hat{U}_j \hat{U}_t$ is the residual of regression of $y_t$ on $D_t$ and $\hat{\lambda}^2$ is estimate of the long run variance of $U_t$ using $\hat{U}_t$.

- **Philips-Perron (PP) Test**

  Philips and Perron (1998) designed more powerful test called PP, which apply to more wider class of error process than AR(p) model. The PP test correct any serial correlation and heteroscedasticity in the disturbance terms.
  The hypothesis $H_0 : \rho = 1$ is based on this model

$$\Delta y_t = \rho y_{t-1} + (constant, timetrend) + \epsilon_t$$

and its test statistic is defined as:

$$Z = (n-1)(p-1) - \frac{\frac{1}{2}(S_{np}^2 - S_{\epsilon}^2)}{\frac{1}{n-1}\sum_{i=1}^{p}(y_{t-1} - \bar{y})}$$

where

$$S_{np}^2 = \sum_{t=2}^{n} \epsilon_{t-1}^2 + \frac{2}{n-1}\sum_{i=1}^{p}\sum_{j=1}^{p} \hat{\epsilon_t}\hat{\epsilon_{t-j}}$$

and

$$S_{\epsilon}^2 = \frac{1}{n}\sum_{t=2}^{n} \epsilon_t^2$$

### 2.0.4 Time Series Models

Box and Jenkins (1976) proposed the Autoregressive Integrated Moving Average (ARIMA) in order to model volatility of financial assets.

- **Autoregressive (AR) Model**

  The AR(p) model is defined as:

  $$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t$$

  where $\mu$ is a constant mean $\phi_i, i = 1, 2, ..., p$ are the parameters of the model and $\epsilon_t$ is white noise.

- **Moving Average (MA) Model**

  The MA(q) model is defined as:

  $$y_t = \mu + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

  where $\theta_i, i = 1, 2, ..., q$ are the parameters of the model

- **Autoregressive Moving Average (ARMA) Model**

  The ARMA(p,q) model is defined as:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

- **Autoregressive Integrated Moving Average (ARIMA)**

  The ARIMA(p,d,q) model is defined as:

$$\phi(B)(1-B)^d X_t = \theta(B) Z_t \quad where \quad Z_t \sim WN(0, \sigma^2)$$

  where $B$ is the backward shift operator such that $BX_t = X_{t-1}$ and the integration parameter $d$ is nonnegative integer.

- **Autoregressive Fractionally Integrated Moving Average (ARFIMA) Model**

  Granger and Joyeux (1980) and Hosking (1981) developed long-memory models to improve the basic ARIMA processes with the pioneering study on the autoregressive fractionally integrated moving average (ARFIMA) model. The ARFIMA (p,d,q) is defined as:

$$\Phi(B)(1-B)^d X_t = \Theta(B) \epsilon_t \quad for \quad d \in (-0.5, 0.5)$$

  where the polynomial $\Phi(B) = 1 - \phi_1 B - ... - \phi_p B^p$ and $\Theta(B) = 1 + \theta_1 + ... + \theta_q B^q$ have order $p$ and $q$ respectively with all their roots outside the unit circle. The $(1-l)^d$ is the fractional difference operator defined by:

$$(1-L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d) L^k}{\Gamma(-d)\Gamma(k+1)}$$

  with $\Gamma(.)$ denoting the gamma function. The process is said to exhibit intermediate long range negative dependence for $d \in (-0.5, 0)$. The process exhibit short memory for $d = 0$, corresponding to stationary and invertible ARIMA model. Then for $d \in [0.5, 1)$ the process is mean reverting. The spectral density function of ARFIMA model is defined as:

$$f_x(\omega) = f_v(\omega)(2Sin(\frac{\omega}{2})), \quad \omega \in [-\pi, \pi]$$

where $f_v(\omega)$ is the spectral density function of the process $(1 - B)^d X_t$.

The ARIMA and ARFIMA model have a weakness assumption of ignoring heteroskedasticity or the presence of unconstant variance, which is an important feature of non-stationarity in predicting time-series data. As result they failed to capture the stylised facts (Cont, 2001) of financial returns such as leptokurtosis, volatility clustering, heavy tails, leverage (asymmetric) effects.

### 2.0.5 Steps in Model building Strategy

Box and Jenkins (1976) developed a three stage procedure for models building:

- **Model Identification**

  Model identification refers to the methodology in identifying required transformation, polynomial orders of $AR(p)$ and $MA(q)$ which are used as a guide in choosing ARMA model that seem to be appropiate. Model identification procedures are as follows:

  Plot the time series data and select the appropriate transformation.

  Estimate the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the original series to further conform the method of transformation and differencing so as that the series is stationary. If the sample ACF decays very slowly and the sample PACF cuts off after lag k, then it indicates that differencing is needed.

  Estimate and examine ACF and PACF of the properly transformed series to identify $p$ and $q$. If the ACF tails off in an exponentially decay from $p$ or contain dumped fluctuations and PACF cuts off after $lag\ \ (p)$ then process is $AR(p)$. If ACF cuts off after $q$ lags and PACF tails off as exponentially or may contains dumped fluctuations, then the process is $MA(q)$. If ACF and PACF both exhibit a decay exponentially from $max(p, q)$ contains dumped fluctuations, then the process is an ARMA.

- **Model Estimation**

  The most used methods for estimating ARIMA models are two step regression, Yule Walker estimation, method of moments and method of maximum likelihood estimation (MLE). However, in recent literature they have widely utilised the MLE method because of its asymptotic properties, consistent, asymptotically distributed and asymptotically efficient.

**Maximum Likelihood Estimation (MLE)**

Let $\Theta = (\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$ denotes the vector of population parameter. Suppose we have observed a sample size T

$$X = (x_1, ..., x_T)$$

Let the joint probability density function (pdf) of these data be denoted

$$f(x_T, x_{T-1}, ..., x_1, \theta)$$

The likelihood function is this joint density treated as a function of the parameter $\theta$ given the data $x$:

$$L(\theta|x) = f(x_T, x_{T-1}, ..., x_1, \theta)$$

The MLE is

$$\hat{\theta} = argmax L(\theta|x)$$

, where $\Theta$ is the parameter space.

- **Model Diagnostic**

  One of the important issues to address before the application of GARCH-type models, the residuals of ARIMA model have to exhibit heteroscedasticity. Engle (1982) prposed the Lagrange Multiplier (LM) test for ARCH effect, to test the presence of heteroscedasticity in returns.

**Lagrange Multiplier (LM) test** Consider an ARMA(p,q) model for the conditional mean equation:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

after obtaining the residuals $\epsilon_t$, the next step is regressing the squared residuals on a constant and $q$ lags as in the following equation:

$$\epsilon_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i} + v_t$$

The null hypothesis that there is no ARCH effect up to order $q$ can be formulated as:

$$H_0 = \alpha_1 = \alpha_2 = ... = \alpha_q$$

against the alternative hypothesis:

$$H_1 = \alpha_i > 0$$

for at least one $i = 1, 2, ..., q$.

The test statistic for the joint significance of the q-lagged square residuals is the number of observations time the R-squared ($TR^2$) from the regression. $TR^2$ is compared against $\chi^2_{(q)}$.

### 2.0.6 Estimation of $d$ for ARFIMA model

There are vast of methods in literature that are used to estimate $d$ . The most widely are four semi-parametric and regression equations constructed from the spectral density function $f_x(\omega)$ and the other one is a parametric method.

- **Periodogram Estimator** $(\hat{d}_p)$

  Geweke and Porter-Hudak (1983), proposed $(\hat{d}_p)$ which uses the periodogram function $I(\omega)$ as an estimate of the spectral density function of $f_x(\omega)$. The number of observation in regression equation is a function $g(n)$ of the sample size $n$ where $g(n) = n^{\alpha}$, $0 < \alpha < 1$

- **Smoothed Periodogram Estimator** $(\hat{d}_{sp})$

  Reisen (1994) developed $(\hat{d}_{sp})$ in the sequel and the regression estimator is obtained by replacing the $f_x(\omega)$ by the smoothed periodogram function with Parzen window. In this procedure $g(n) = n^{\alpha}$, $0 < \alpha < 1$ and the truncation point in the Parzen lag window is $m = n^{\beta}$, $0 < \beta < 1$

- **Robinson Estimator** $(\hat{d}_{pr})$

  Robinson (1995) introduced $(\hat{d}_{pr})$, which regresses $\{InI(\omega_i)\}$ on $In(2Sin(\frac{\omega_i}{2}))^2$ for $i = I, I+1, ..., g(n)$, where $I$ is the lower truncation point tents to infinity more slowly than $g(n)$.

- **Robinson Estimator based on Smoothed Periodogram** $(\hat{d}_{spr})$

  Robinson (1995) extended the Robinson estimator by the use of the smoothed periodogram function, with the Parzen lag window. The truncation point is $g(n) = n^{\alpha}$, $0 < \alpha < 1$ and the number of observation in the regression equation is the same with one chosen by $\hat{d_{pr}}$.

- **Whittle Estimator** $(\hat{d}_{\omega})$

  Whittle (1953) developed $(\hat{d}_{\omega})$ which a parametric method and Fox and Taqqu (1986) suggested modification for $(\hat{d}_{\omega})$. The $(\hat{d}_{\omega})$ is based on the periodogram and involves the function

$$Q(\zeta) = \int_{\pi}^{\pi} \frac{I(\omega)}{f_x(\omega, \zeta)} d\omega$$

where $f_x(\omega, \zeta)$ in known spectral density function at frequency $\omega$ and $\zeta$ denotes the vector of unknown parameters. The $(\hat{d}_\omega)$ is the value of $\zeta$ which minimises the function $Q(.)$.

### 2.0.7 Autoregressive Conditional Heteroscedasticity (ARCH)

In order to address the empirical observation that variance is heteroscedasticity with time and that it seems to be dependent on its past values. Engle (1982) proposed the Autoregressive Conditional Heroscedasticity (ARCH) to model volatility by relating the conditional variance of the disturbance terms to a linear combination of squared disturbances in the recent past. The conditional variance of the error terms $\epsilon_t$, is modeled as being normally distributed with mean zero and constant variance $\sigma_t^2$:

$$\sigma_t^2 = \omega + \sum_{i=1}^{m} \alpha_i \epsilon_{t-i}^2$$

where $\alpha_i > 0$ for $i = 0, 1, ...m$ since the variance cannot be non negative. However, the ARCH model could not capture the following stylised facts such as volatility clustering, leverage effects, long range dependence and structural breaks.

### 2.0.8 Symmetric GARCH-type models

The GARCH, IGARCH and GARCH-M are considered to be symmetric GARCH-type models, which imply that the positive and negative shocks is equal the size exhibited by an equal response from the market. Symmetric GARCH-type models can capture three important stylsed facts of financial time series .i.e leptokurtosis, skewness and volatility clustering. However, they cannot capture dynamics of leverage effects, long range dependence and structural breaks.

- **Generalised ARCH (GARCH)**

Bollerslev(1986) proposed Generalised ARCH (GARCH) model in which the conditional variance depends simultaneously on the squared residuals and its own past values.

mean equation:

$$r_t = \mu + \epsilon_t$$

conditional variance equation:

$$\sigma_t^2 = \omega + \sum_{i=1}^{r} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{m} \beta_i \sigma_{t-i}^2$$

where $r_t$ is the return of an asset at time t, $\mu$ is the average returns, $\epsilon_t$ is the residual returns, $\omega > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$ and $\sigma_t$ is the conditional variance $r$ in the order of the Moving Average (MA) ARCH term and $m$ in the order of Autoregressive (AR) GARCH term. The size of parameter $\alpha$ and $\beta$ determine the short run dynamics of volatility time series. However, GARCH (r,m) can be expressed as ARMA (p,q) process:

$$[1 - \alpha(L) - \beta(L)]\epsilon_t^2 = \omega + [1 - \beta(L)]v_t$$

where $v_t = \epsilon_t^2 - \sigma_t^2$, $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + ... + \alpha_r L^r$ and $\beta(L) = \beta_1 L + \beta_2 L^2 + ... + \beta_m L^m$. If the sum of the coefficient is equal to one ($\alpha + \beta = 1$) then any shock will lead to a permanent change in all future values of volatility. Hence this leads to a persistent shock to the conditional variance.

- **Integrated GARCH (IGARCH)**

Engle and Bollerslev (1986) proposed the Integrated GARCH (IGARCH) to accommodate persistent conditional volatility. If the sum of the coefficient equals 1 of GARCH it is called an IGARCH.
Conditional variance model:

$$\phi(L)(1-L)\epsilon^2 = \omega + (1-\beta(B))v_t$$

where $\phi(L) = \sum_{i=1}^{m} \phi_i L^i$ and $1-L$ is the first difference operator.

- **GARCH-in-Mean (GARCH-M)**

Engle Lilien and Robins (1987) introduced GARCH-in -Mean (GARCH-M) model where returns of security may depend on its volatility. The GARCH-M was proposed for modeling risk return tradeoffs. In financial investment theory it is predicted that the expected returns on an asset is directly proportional to the expected risk of an asset. This implies that, high risk is often expected to lead to high returns as a compensation for taking risk.
mean equation:

$$r_t = \mu + \lambda\sigma_t + \epsilon_t$$

conditional variance equation:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{m} \beta_i \sigma_{t-i}^2$$

The parameter $\lambda$ in the mean equation is called the risk premium. A positive value of $\lambda$ signifies that the return is positively related to its volatility. i.e. a rise in mean return is caused by increase in conditional variance as a proxy of increase in risk. The time varying risk premium is estimated by the significance of $\lambda$ coefficient of $\sigma_t$ in the mean equation.

### 2.0.9 Asymmetric GARCH-type models

Enders (2004) the tendency for volatility to decline when returns rise and rise when returns fall is often called the leverage effect.

- **Exponential GARCH (EGARCH)**

Nelson (1991) proposed an Exponential GARCH (EGARCH) model, which is the logarithmic expression of the conditional volatility used to capture the asymmetric effects. The general form of the EGARCH (r,m)

$$ln\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r} \beta_i ln\sigma_{t-i}^2 + \sum_{i=1}^{m} \{\alpha_i |\frac{\epsilon_{t-1}}{\sigma_{t-i}}| - \sqrt{\frac{2}{\pi}}\} - \gamma \frac{\epsilon_{t-i}}{\sigma_{t-i}}$$

where $\gamma$ is the asymmetric or leverage effect. The sign $\gamma$ is expected to be positive so that a negative shock increases future volatility or uncertainty which a positive shock eases the effect on future uncertainty.

- **Power GARCH (PGARCH)**

Ding, Granger and Engle (1993) introduced the standard deviation GARCH model, whereby the standard deviation is model led rather than the variance and is called Power GARCH (PGARCH). In PGARCH an optional parameter $\gamma$ account for the leverage effect. The PGARCH(r,m) is expressed as:

$$\sigma_t^{\delta} = \alpha_0 + \sum_{i=1}^{m} \beta_i \sigma_{t-i}^{\delta} + \sum_{i=1}^{r} \alpha_i (|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i})^{\delta}$$

when $\delta$=2 becomes a classical GARCH model that captures the leverage effects. However, when $\delta$=1 the conditional standard deviation will be estimated.

- **Treshold GARCH (TGARCH)**

Zakoian (1994), proposed Threshold GARCH (T-GARCH) model which allows analysis on the effect of good news and bad news (negative and positive returns) on volatility. It is also known as the GJR-GARCH model which was proposed by Glosten, Jagannathan, and Runkle in 1993. TGARCH (r,m) the conditional variance equation is specified as:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r}(\alpha_i + \gamma_i I_{t-i})\epsilon_{t-i}^2 + \sum_{i=1}^{m}\beta_i\sigma_{t-i}^2$$

where $I_{t-i}$ is the dummy variable used to distinguish the good news and bad news i.e $I_{t-1}$=1 if $\epsilon_{t-i}$ <0 indicating bad news and if $I_{t-i}$=0 if $\epsilon_{t-i}$ ≥0 indicating good news. Good news has an impact of $\alpha_i$ while bad news has an impact of $\alpha_i + \gamma$. Ahmed and Saliman (2011), a non-zero value of $\lambda$ indicates the asymmetric of the returns.

### 2.0.10 Long Memory GARCH-type models

The long-memory characteristic of time-series data offers explanation on the persistent temporal dependence between distant observations, which implies their predictability. The evolution of forecasting in time-series data using the autoregressive (AR) methodology started when Box and Jenkins (1976) popularised the general autoregressive moving average (ARMA) function. The combined AR and MA models have advantages in predicting time-series data because of the effective yet simple and explicit model structures focusing on the mean.

- **Fractionally Integrated GARCH (GARCH)**

Baillie, Bollerslev and Mikkelsen (1996) proposed the Fractionally Integrated GARCH (FIGARCH) model to accommodate long range dependence and persistence in the conditional volatility. The FIGARCH is obtained by replacing the first difference operator of IGARCH by $(1-L)^d$ where $d$ is a fractional $0 < d < 1$. Thus the FIGARCH model can be expressed as:

$$\epsilon_t^2 = \omega + [1 - B(L)][(1-L)^d \phi(L)]^{-1} v_t$$

- **Fractionally Integrated Exponential GARCH (FIEGARCH)**

Bollerslev and Mikkelsen (1996) extended the FIGARCH model to Fractionally Integrated Exponential GARCH (FIEGARCH) to capture volatility clustering, long memory and leverage effect. The extension of exponential GARCH the FIEGARCH defined as:

$$In(\sigma_t^2) = \alpha_0 + \frac{1 - \sum_{i=1}^{r} \alpha_i L^i}{1 - \sum_{j=1}^{m} \beta_j L^j} (1-l)^{-d} g(\epsilon_{t-1})$$

where $g(\epsilon_{t-1}) = \theta_{\epsilon_{t-1}} + \gamma[|\epsilon_{t-1}| - E(|\epsilon_{t-1}|)]$

- **Adaptive Fractionally Integrated GARCH (AFIGARCH)**

Baille and Moran (2009) proposed an Adaptive FIGARCH (AFIGARCH) model which accommodates long range dependence and structural changes in the volatility.

$$[1 - B(L)]\sigma_t^2 = \omega + [1 - B(L) - \phi(L)(1 - L)^d]y_t^2$$

where $y_t$ is a discrete time, real valued stochastic process that is serially uncorrelated in its conditionally mean and has long range dependence type in its conditional variance process.

- **Smooth Transition FIGARCH (STFIGARCH)**

Kolic (2007) introduced a smooth transition FIGARCH (STFIGARCH) to accommodate smooth changes both in the amplitude of volatility clustering, leverage effects and long range dependence in the conditional volatility.

$$\phi(L)(1 - L)^d\epsilon_t^2 = \omega + B(L)v_t$$

- **Fractionally Integrated Time Varying GARCH (FITVGARCH)**

Ben Nasr et al. (2010) introduced the fractionally integrated time varying GARCH (FITVGARCH) model to capture long memory, structural changes in the volatility and allows the conditional volatility to change over time.

$$[1 - \phi_t(L)](1 - L)^d\epsilon_t^2 = \omega + [1 - B_t(L)]v_t$$

### 2.0.11 Model Selection

Model selection among the GARCH-type models is performed using various criteria:

- **Akaike Information Criteria (AIC)** due to Akaike (1974) defined by:

$$AIC = 2k - 2InL(\hat{\Theta})$$

  where k denotes the number of unknown parameters, $\Theta$ the vector of the unknown parameter and $\hat{\Theta}$ the maximum likelihood estimate.

- **Bayesian Information Criteria (BIC)** due to Schwarz (1978) define by:

$$BIC = kInn - 2InL(\hat{\Theta})$$

  where n denotes the number of observations.

- **Hannan-Quinn Criteria(HQC)** due to Hannan and Quinn (1979) defined by:

$$HQC = -2InL(\hat{\Theta}) + 2kInInn$$

- **Consistent Akaike Information Criteria (CAIC)** due to Bozdogan (1987) defined by:

$$CAIC = -2InL(\hat{\Theta}) + k(Inn + 1)$$

- **Corrected Akaike Information Criteria (AICc)** due to Hurvich and Tsai (1989) defined by:

$$AICc = 2k - 2InL(\hat{\Theta}) + \frac{2k(k+1)}{n-k-1}$$

**NB** The smaller the values of AIC, BIC, HQC, CAIC and AICc better the fitted model.

### 2.0.12 Value at Risk (VaR)

Christoffersen (2003) defines Value at Risk (VaR) as risk concept developed in the area of portfolio risk management which summarises the downside risk in a single statistic. VaR can be defined as the least amount of money that one could expect to lose from an investment given probability over a specific period of time. Financial institutions can have a sense on the minimum amount that is expected to lose with a small probability $\alpha$ over a given time k (usually 1 day or 10days).

Let $P_t$ be the price of a financial asset on a day k. A k-day VaR on day t is defined by:

$$P(P_{t-k} - P_t \leq Var(t,k,\alpha)) = 1 - \alpha$$

Jorion (2001) given a distribution of returns, VaR can be derived and expressed in terms of percentiles of the return distribution.

$$VaR(t,k,\alpha) = (1 - \exp^{q_\alpha})P_{t-k}$$

where $\alpha^{th}$ percentile of the returns. When using GARCH-type models, VaR is calculated directly from the standard deviation obtained from fitting the distribution of returns.

### 2.0.13 Estimation VaR

The estimation of one-day ahead VaR considering eight widely used innovation distributions and GARCH-type models.

- The normal distribution:

  $$V\hat{a}R_p = \mu_{\hat{t+1}} + \sigma_{\hat{t+1}}\Phi^{-\hat{1}}(p),$$

  where $\Phi(.)$ is the distribution function of the standard normal distribution.

- Azzaline (1985) defined VaR for skewed normal distribution:

$$\hat{VaR_p} = \hat{\mu_{t+1}} + \hat{\sigma_{t+1}}\hat{\Phi^{-1}}(p),$$

where the distribution function:

$$F(y) = \Phi(x) - 2T(y, \hat{\alpha})$$

for $-\infty < y < \infty$ and $-\infty < \alpha < \infty$, where $T(.,.)$ denotes Owen's T function (Owen (1956)).

- Gosset (1908) defined VaR for Student's t distribution:

$$\hat{VaR_p} = \hat{\mu_{t+1}} + \hat{\sigma_{t+1}}\hat{\Phi^{-1}}(p),$$

where the distribution function:

$$F(y) = \int_{-\infty}^{y} f(w)dw$$

for $-\infty < y < \infty$, where probability density function:

$$f(y) = \frac{2\Gamma(\frac{\hat{v}+1}{2})}{\sqrt{\hat{v}(\hat{v}-2)\Gamma(\frac{\hat{v}}{2})}}(1 + \frac{y^2}{\hat{v}-2})^{-\frac{\hat{v}+1}{2}}$$

for $-\infty < y < \infty$ and $\hat{v} > 0$

- Fernandez and Steel (1998) defined VaR for skewed Student's distribution:

$$\hat{VaR_p} = \hat{\mu_{t+1}} + \hat{\sigma_{t+1}}\hat{\Phi^{-1}}(p),$$

where the distribution function:

$$F(y) = \int_{-\infty}^{y} f(w)dw$$

for $-\infty < y < \infty$, where probability density function:

$$f(y) = \frac{2\Gamma(\frac{\hat{v}+1}{2})}{\sqrt{\hat{v}(\hat{v}-2)}(\hat{\xi} + \hat{\xi}^{-1})\Gamma(\frac{\hat{v}}{2})} \begin{cases} (1 + \frac{\hat{\xi}^2 y^2}{\hat{v}-2})^{-\frac{\hat{v}+1}{2}} & \text{if} \quad y < 0 \\ (1 + \frac{y^2}{\hat{\xi}(\hat{v}-2)})^{-\frac{\hat{v}+1}{2}}, & \text{if} \quad y \geq 0 \end{cases}$$

where $-\infty < y < \infty$, $v > 0$ and $\xi > 0$

- Theodossiu (1998) defined VaR for the skewed generalized error distribution:

$$V\hat{a}R_p = \hat{\mu_{t+1}} + \hat{\sigma_{t+1}}\hat{\Phi^{-1}}(p),$$

where the distribution function:

$$F(y) = \int_{-\infty}^{y} f(w)dw$$

for $-\infty < y < \infty$, where probability density function:

$$f(y) = \begin{cases} 1 + [\frac{\hat{k}\hat{\theta}^{-\hat{k}}(1-\hat{\lambda})^{-\hat{k}}}{\hat{n}-2}|x|^{\hat{k}}]^{-\frac{\hat{n}+1}{\hat{k}}} & \text{if} \quad y < 0 \\ 1 + [\frac{\hat{k}\hat{\theta}^{-\hat{k}}(1+\hat{\lambda})^{-\hat{k}}}{\hat{n}-2}|x|^{\hat{k}}]^{-\frac{\hat{n}+1}{\hat{k}}}, & \text{if} \quad y \geq 0 \end{cases}$$

for $-\infty < y < \infty$, $k > 0$, $n > 2$ and $-1 < \lambda < 1$.

- Bandorff-Nielsen (1977, 1978) defined the generalized hyperbolic distribution:

$$V\hat{a}R_p = \hat{\mu_{t+1}} + \hat{\sigma_{t+1}}\hat{\Phi^{-1}}(p),$$

where the distribution function:

$$F(y) = \int_{-\infty}^{y} f(w)dw$$

for $-\infty < y < \infty$, where probability density function:

$$f(y) = C\sqrt{1+y^2}^{\hat{\lambda}-\frac{1}{2}} K_{\hat{\lambda}-\frac{1}{2}}(\hat{\alpha}\sqrt{1+y^2}) \exp{(\hat{\beta}x)}$$

for $-\infty < y < \infty$, $-\infty < \lambda < \infty$, $-\infty < \alpha < \infty$, $-\infty < \beta < \infty$, $|\beta| < \alpha$ if $\lambda > 0$, $|\beta| < \alpha$ if $\lambda = 0$ and $|\beta| \leq \alpha$ if $\lambda < 0$ where $C$ is a normalised constant and $K_v(.)$ denotes the modified Bessel function of the second kind defined by:

$$K_v(y) = \begin{cases} \frac{\pi csc(\pi v)}{2}[I_{-v}(y) - I_v(y)] & \text{if} \quad v \notin 0 \\ \lim\limits_{\mu v} K_\mu(y) & \text{if} \quad v \in 0 \end{cases}$$

where $I_v(.)$ denotes the modified Bessel function of the first kind of order $v$ defined as:

$$I_v)(y) = \sum_{k=0}^{\infty} \frac{1}{\Gamma(k+v+1)k!}(\frac{y}{2})^{2k+v}$$

- Barndorff Nielsen (1977, 1978) defined VaR for the normal inverse Gaussian distribution:

$$V\hat{a}R_p = \hat{\mu_{t+1}} + \hat{\sigma_{t+1}}\Phi^{-\hat{1}}(p),$$

where the distribution function:

$$F(y) = \int_{-\infty}^{y} f(w)dw$$

for $-\infty < y < \infty$, where probability density function:

$$f(y) = C\sqrt{1+y^2}^{\hat{\lambda}-\frac{1}{2}} K_1(\hat{\alpha}\sqrt{1+y^2}) \exp{(\hat{\beta}x)}$$

for $-\infty < y < \infty$, $-\infty < \alpha < \infty$, $-\infty < \beta < \infty$ and $|\beta| \leq \alpha$ where C is a normalizing constant.

- Johnson (1949) defined the Johnson's SU distribution:

$$V\hat{a}R_p = \hat{\mu_{t+1}} + \hat{\sigma_{t+1}}\Phi^{-\hat{1}}(p),$$

where the distribution function:

$$F(y) = \Phi(\hat{\alpha} + \hat{\beta} arcsinh(y))$$

for $-\infty < y < \infty$, $\alpha > 0$ and $\beta > 0$.

### 2.0.14 Backtesting VaR

In order to measure the accuracy of estimated VaR, the risk models can be Backtested using the Unconditional coverage test, Conditional coverage test, Loss functions and Dynamic Quantile tests.

**Kupiec Test**

Kupiec (1997) proposed a Likelihood Ratio ($LR$) test of unconditional coverage ($LR_{uc}$) which accounts only for the frequency of VaR violations and evaluates the model accuracy. The Kupiec test examines whether the failure rate is equal to the expected value. The null hypothesis

$$H_0 : p = \hat{\pi}$$

The LR test statistic is given by:

$$LR_{uc} = -2In\left[\frac{p^{n_1}(1-p)^{n_0}}{\pi^{n_1}(1-\pi)^{n_0}}\right] \sim \chi^2_{(1)}$$

where $\pi = \frac{n_1}{n_0+n_1}$ is the MLE of $p$, $n_1$ represent the total violations and $n_0$ the total non violations forecast.

**Christoffersen Test**

Christoffersen (1998) developed a Likelihood Ratio (LR) test of conditional coverage $LR_{cc}$ which accounts for both the frequency of VaR violations and the time at which they occur and also evaluates the model accuracy. The LR test statistic is given by:

$$LR_{cc} = -2In\left[\frac{(1-p)^{n_0}p^{n_1}}{(1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}^{n_{11}}}\right] \sim \chi^2_{(2)}$$

where $\pi_{01} = \pi_{11} = p$ and $nij$ is the number of observation with value $i$ followed by $j$ for $i, j = 0, 1$ and $\pi_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}$

Advantages of Christoffersen test over Kupiec test:

- Accounts for conditionality in forecasting
- Enables us to separate clustering effects from assumption effects

- It can reject a VaR model that generates either too many or few clustered violation

**Loss Fuctions**

Sarma et al (2003) proposed three test, Regulator Loss Function (RLF), Unexpected Loss Function (ULF) and Firm's Loss Function (FLF) to account for some models that may have same violation number with different errors.

- RLF test to take into account different realised returns and VaR forecasts.

$$RLF_{t+1} = \begin{cases} (r_{t+1} - VaR_{t+1})^2, & \text{if} \quad r_{t+1} < VaR_{t+1} \\ 0, & \text{if} \quad r_{t+1} \geq VaR_{t+1} \end{cases}$$

  $VaR_{t+1}$ represents the one-day-ahead VaR forecast for a long position.

- ULF test is the same as the average value of differences between realised returns and VaR forecasts. realised returns and VaR forecasts.

$$ULF_{t+1} = \begin{cases} r_{t+1} - VaR_{t+1}, & \text{if} \quad r_{t+1} < VaR_{t+1} \\ 0, & \text{if} \quad r_{t+1} \geq VaR_{t+1} \end{cases}$$

- FLF considers the situation in which the realised returns exceeds the VaR forecast. realised returns and VaR forecasts.

$$FLF_{t+1} = \begin{cases} (r_{t+1} - VaR_{t+1})^2, & \text{if} \quad r_{t+1} < VaR_{t+1} \\ \beta VaR_{t+1}, & \text{if} \quad r_{t+1} \geq VaR_{t+1} \end{cases}$$

  where $\beta$ is the cost excess capital.

**Dynamic Quantile (DQ) test**

Engle and Manganelli (2004) proposed Dynamic Quantile (DQ) test, which examines if violations is uncorrelated with any other variables, when VaR is computed. DQ test regress the current violations in order to test for different restrictions on the parameters of the model. The linear regression model:

$$I_t = \beta_0 + \sum_{i=1}^{p} \beta_i I_{t-i} + \sum_{j=1}^{q} \mu_j X_j + \epsilon_t$$

where

$$I_t = \begin{cases} 1, & \text{if} \quad r_{t+1} < VaR_{t+1} \\ 0, & \text{if} \quad r_{t+1} \geq VaR_{t+1} \end{cases}$$

### 2.0.15 Evaluation of Forecasting Performance

Evaluating the performance of different forecasting models plays a critical role in choosing the most accurate models.

- Logarithmic Loss (LL) due to Pagan and Schwert (1990) defined by

$$LL = \sqrt{\frac{1}{N-n}\sum_{t=n+1}^{N}(log(y_t^2) - log(\hat{\sigma_t}))^2}$$

  where $y_t is$ is the observation at time $t$ and $\hat{\sigma_t}$ the predicted conditional volatility at time $t$, n is the total observations in the in-sample period and N is the total number of observations.

- Heteroskedasticity adjusted Mean Square Error (HMSE) due to Bollerslev and Ghysels (1996) defined by

$$HMSE = \sqrt{\frac{1}{N-n}\sum_{t=n+1}^{N}(\frac{y_t^2}{\hat{\sigma_t}} - 1)}$$

- Theil Inequality (TI) due to Theil (1996) defined by

$$TIC = \frac{\sqrt{\frac{1}{n}\sum_{1}^{n}(\sigma_t - \sigma_t)^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}\sigma_t} + \sqrt{\frac{1}{n}\sum_{t=1}^{n}\sigma_t}}$$

  where $\sigma_t$ is observed conditional volatility.

Other methods that are used to evaluate forecasting performance of GARCH-type models, error functions are used. Mostly used functions in literature are:

- Mean Error (ME)

$$ME = \frac{1}{N}\sum_{t=1}^{N}(\hat{\sigma}_t - \sigma_t)$$

- Mean Square Error (MSE)

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(\sigma_t - \sigma_t)^2$$

- Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\sigma_t - \sigma_t)^2}$$

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|\hat{\sigma}_t - \sigma_t|$$

- Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{|\hat{\sigma}_t - \sigma_t|}{\sigma_t}$$

# Chapter 3

# METHODOLOGY

## 3.1 Data Description

The chapter discusses methods and techniques that will be adopted in analyses of the data. The time series data will be used for modelling financial stylized facts of British Petroleum returns. The daily British Petroleum returns will be adopted from 01 January 1998 to 31 December 2018 resulting in 5282 observation excluding weekends and holidays. British Petroleum returns will be computed as the continuously compound returns which were the first difference in logarithm.

$$r_t = log\frac{P_t}{P_{t-1}} \tag{3.1.1}$$

where $P_t$ and $P_{t-1}$ are British Petroleum closing prices at current day and previous day respectively.

## 3.2 Time Series Plots

Time series data is a sequence of points, typically measured at successive time instants spaced at uniform time interval. In financial time series data is used for identifying stationarity, volatility clustering and leverage effects. In this study time series plot for British Petroleum prices and compound continuously returns will be plotted.

## 3.3 Summary Statistics

The summary statistics will be computed to inspect characteristics of returns such as mean, median, standard deviation, skewness, kurtosis. Standard deviation is a measure of volatility of financial time series data. Skewness is the measure of how asymmetric returns relative to the Gaussian distribution. If the skewness is positive, it suggests that the distributions of returns has a long left tail. If the skewness is negative, it suggests that the distribution of returns has long right tail. Kurtosis is the measure of peakedness of returns. Kurtosis of the Gaussian distribution is 3. If the kurtosis exceeds 3, then distribution is peaked (leptokurtic) relative to Gaussian distribution. If the kurtosis is less than 3, then distribution is flat (platykurtic) relative to the Gaussian distribution.

## 3.4 Diagnosis Tests

### 3.4.1 Stationarity Test

One of the fundamental concepts underpinning financial time series analysis is stationarity. The widely used test for stationarity is Phillips-Perron because it is more robust in the presence of serial correlation and heteroskedasticity. However, Augmented Dickey Fuller is to be conducted to confirm the same results with Phillips-Perron test for unit root.

### 3.4.2 Normality Test

The test of normality can be assessed by constructing Q-Q plots of returns, conducting Jarque-Bera (JB) test and Darling Anderson test. Q-Q plot is a scatter plot of the empirical quantiles (vertical axis) against the theoretical quantiles (horizontal axis) of a given distribution. If the returns follow approximately a normal distribution with constant mean and variance, then the resulting of the plot should be roughly scattered around the 45 degree line with positive slope. The JB test assumes that data follow a normal distibution and it is distributed as chi-square with 2 degrees of freedom.

### 3.4.3 Heteroscedasticity Test

It is intuitive to assesses if there are ARCH effects in the residuals of the mean model before we fit the GARCH-type model. This study focuses on the Engle's Langrage Multiplier in testing Heteroskedasticity.

### 3.4.4 Autocorrelation Test

GARCH-type models to be applied require that returns are serially uncorrelated, but still dependent (Tsay, 2005,99). In financial time series modelling it is of paramount importance to test jointly that several autocorrelation of returns are zero. Ljung and Box test will be conducted because it enhance the power of test in finite sample.

### 3.4.5 Structural Break Test

In modelling time series data it is key to check for structural breaks in returns. Chow test and CUSUM test based on ordinary least squares will employed to asses structural breaks of returns.

### 3.4.6 Long Memory Test

In order to model long memory GARCH-type models, squared returns will be tested for the presence of long range dependence by the following test periodogram estimator (GPH) and wavelet estimator which are parametric and semi-parametric methods respectively.

## 3.5 Heavy Tailed Distribution and Selection

The fat tails of returns are to be captured by eight distributions of innovation namely; normal distribution, skewed normal distribution, student-t distribution, skewed student-t distribution, generalized error distribution, skewed generalized error distribution, normal inverse Gaussian distribution and Johnson's SU distribution as distributions of innovations. In financial modelling it is often essential to identify the best distribution of innovation that explains heavy tailed returns. The study will employ the most common selection criteria employed are Akaike information criteria (AIC) and Bayesian information criteria (BIC).

## 3.6 Symmetric GARCH-type models

The symmetric GARCH-type models capture clustering volatility for financial time series data. The three most common GARCH-type models that capture the volatility clustering is GARCH, IGARCH and GARCH -M. The symmetric GARCH-types of model will be estimated assuming distributions of innovation.

## 3.7 Asymmetric GARCH-type models

The asymmetric GARCH-type models captures volatility clustering and leverage effects for the returns. In literature there are vast number of asymmetric GARCH-type models, this study will estimate the following EGARCH, TGARCH and PGARCH assuming distributions of innovation.

## 3.8 Long Memory GARCH-type models

The modeling of long range dependence of returns require specialised GARCH-type models that capture long memory and at the same type capturing volatility clustering and leverage effects. The long memory GARCH-type of models that will be estimated namely; FIGARCH, FIEGARCH and HYGARCH assuming distributions of innovation.

## 3.9 Forecasting

After estimating the different GARCH-type model, the models are used to forecast 1-head daily VaR at both 95 and 99 percent confidence levels.

### 3.9.1 Value at Risk (VaR) Evaluation

There is need to accurately estimate VaR, because the inaccurate estimation of this Value at Risk can have catastrophic repercussion. When using GARCH-type model, VaR is calculated directly from the standard deviation obtained from fitting the distribution of returns.

### 3.9.2 Backtesting VaR

In order to measure the accuracy of estimated risk, the risk models will be Back-tested .Two popular statistics will be used , firstly the unconditional coverage test, which takes into account only the frequency of VaR violations and does not consider the time at which they occur and secondly the conditional coverage test which take both considerations. The most widely used unconditional coverage test is Kupiec test which will be adopted in this study and for conditional coverage test, Christoffersent test will be used to examine the appropriate estimation of VaR.

## 3.10 Evaluation Forecasting Methods

Evaluating the performance of different forecasting models plays a critical role in choosing the most accurate models. In literature are vast number of forecasting evaluation measures, however this study will focus on Mean Absolute Error (MAE), Mean Square Error (MSE), and Root Mean Square Error (RMSE).

# Chapter 4

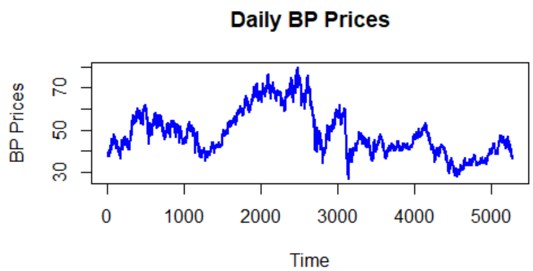
# ANALYSIS AND RESULTS

## 4.1 Introduction

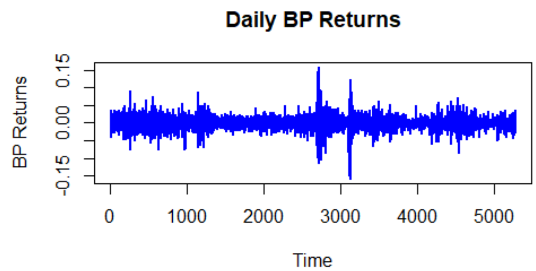
The analysis focused on the daily closing continuously compounded returns of the British Petroleum of London Stock Exchange. The sample data begins from 01 January 1998 to 31 December 2018 excluding weekends and holidays, giving a total of 5282 data points. The financial stylized facts analysis herein includes fitting symmetric, asymmetric and long memory GARCH-type models.

## 4.2 Time Series Plots

Figure 4.1 shows a time series plot of British Petroleum prices and returns. Figure 4.1 (a) exhibits the series of British Petroleum prices showing an expected non-stationary process that exist due to non constant means. In modelling financial time series it must be stationary, hence the closing prices was transformed to continuous compound returns. Figure 4.1 (b) depicts that returns vary around a constant mean but the variance is changing over time, which may suggest the presence of heteroskedasticity. Figure 4.1 (b) infer that high volatility tends to be followed by high volatility for prolonged periods and low volatility tends to be followed by low volatility for prolonged periods, a phenomenon referred to as volatility clustering.




(a)




(b)

Figure 4.1: Time Series Plots

It indicates that GARCH-type models may be appropriate model to explaining British Petroleum returns, since they are constant in means expect clearly non constant variance.

## 4.3 Summary Statistics

Table 4.1 shows the descriptive statistic of British Petroleum returns. The mean of returns is positive, suggesting that British Petroleum prices has increased over the period of 1998 to 2018.

Table 4.1: Descriptive Statistics of Returns

| Statistic | Returns |
|---|---|
| Mean | 0.000152 |
| Standard Deviation | 0.017782 |
| Skewness | -0.065949 |
| Kurtosis | 8.453399 |

The standard deviation is high, indicating a high level of fluctuations of returns. British Petroleum returns show negative skewness, suggesting that the distribution of returns have long left tails. The excess value of kurtosis indicates heavy tailed returns which means the distribution of returns is peaked (leptokurtic) relative to the Gaussian distribution.

## 4.4 Stationarity Tests

In order to check whether the British Petroleum returns are stationary or not, Phillips-Perron test and Augmented Dickey Fuller were conducted and the results are presented in Table 4.2.

Table 4.2: Stationarity Tests of Returns

| Test | Lags | P value |
|---|---|---|
| Philips Perron | 10 | 0.01** |
| Augmented Dickey Fuller | 10 | 0.01** |

*,** and *** represents the significance level at 10, 5, 1 % levels respectively

To test for unit root, Phillips-Perron test in Table 4.2 shows that British Petroleum returns series are stationary in mean at 5% level of significance. We used the Phillips-Perron test since it is robust to the presence of serial correlation and heteroskedasticity. However, Augmented Dickey-Fuller test for unit root was also used to confirm that British Petroleum returns are stationary in mean at 5% level of significance.

## 4.5 Normality Tests

The Q-Q plot of British Petroleum returns shown in Figure 4.2 exhibit the presence of heavy tailed more than Gaussian distribution. This is an indication that British Petroleum returns are not normal distribution. Returns are not normal distributed because of the presence of volatility clustering. Hence modelling returns using GARCH-type models must consider distributions of innovation that capture the fat tails of

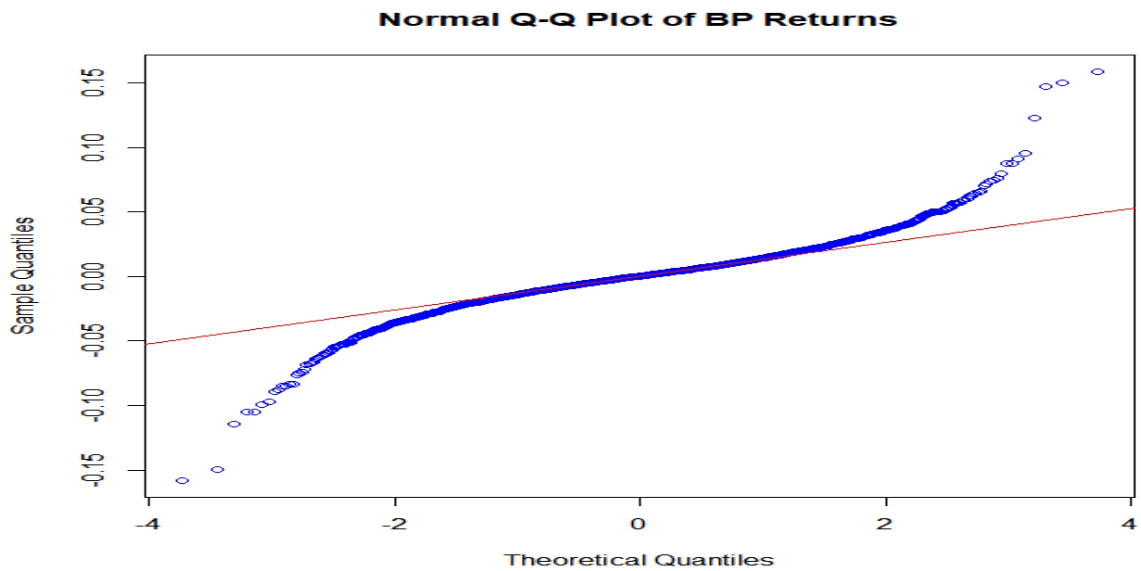



Figure 4.2: Plot of returns

returns. In this study, eight distribution of innovation that have been widely used in literature namely; namely normal distribution, skewed normal distribution, student-t distribution, skewed student-t distribution, generalized error distribution, skewed generalized error distribution, normal inverse Gaussian distribution and Johnson's SU distribution.

Table 4.3 shows Jarque-Bera and Anderson-Darling statistic that have been enormous used for testing normality of returns. In both test, the null hypothesis is rejected at all levels of significance, which illustrates that the British Petroleum returns are not normally distributed.

Table 4.3: Normality Tests of Returns

| Test | Test Statistic | P value |
|---|---|---|
| Jarque-Bera | 15747 | 2.2e-16 |
| Anderson-Darling | 5249.8 | 2.2e-16 |

*,** and *** represents the significance level at 10, 5, 1 % levels respectively

## 4.6 Heteroskedasticity Test

The linear structural models assumes that the variance of errors is constant over time. But this assumption is not applicable for financial time series particularly the stock returns in which the errors exhibit time-varying heteroskedasticity. Before proceeding to apply GARCH-type models, it is necessary to ascertain the existence of ARCH effect in the conditional variance. Figure 4.3 shows the plot of auto-covariance function of squared returns,......

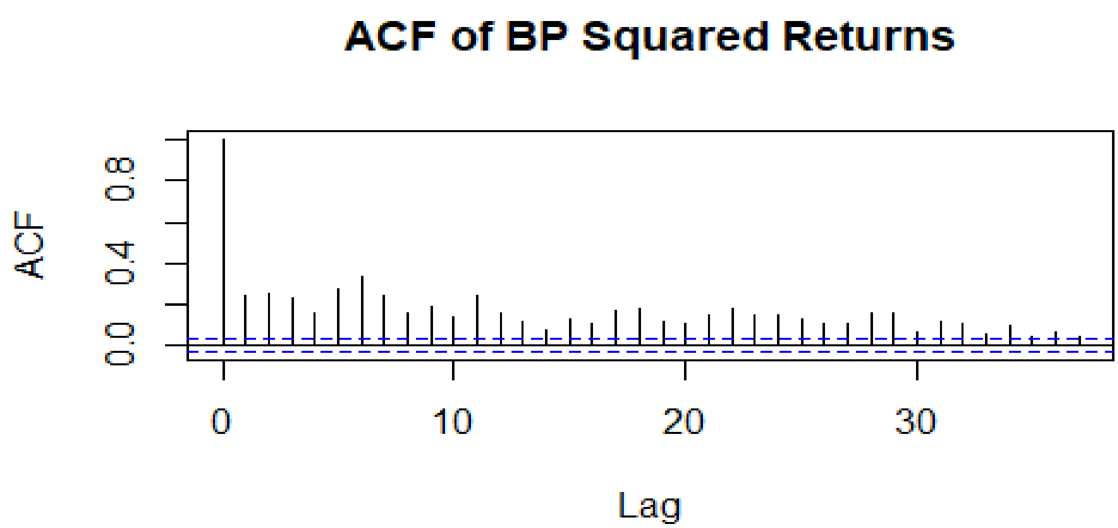



Figure 4.3: ACF of squared returns

We first employed AR(1) model for conditional mean in return series. The null

Table 4.4: Heteroskedasticity Tests of Returns

| Test | Lags | P value |
|---|---|---|
| Langrage Multiplier (LM) | 4 | 0.000*** |
| | 12 | 0.000*** |
| | 24 | 0.000*** |

*,** and *** represents the significance level at 10, 5, 1 % levels respectively

hypothesis that there are no ARCH effects in the residuals series up to lag 24. The ARCH-LM test results in Table 4.4 provide strong evidence for rejecting the null hypothesis, this is an indication of existence of ARCH effect in the residuals series of mean equation and therefore variance of returns are non-constant. The ARCH-LM test confirm the presence of heteroskedaticity which in turn suggest the usefulness and suitability of GARCH-type models for modeling conditional volatility.

## 4.7 Autocorrelation Test

Figure 4.3 shows the autocovariance function of British Petroleum returns becomes significant at most of lag levels except at lag 9, 13, 19, 26, 27. Thus, the correlation among British Petroleum returns are significant.

Table 4.5: Autocorrelation Tests of Returns

| Test | Lags | P value |
|---|---|---|
| Ljung-Box | 10 | 6.084e-06 |
| | 20 | 1.41e-07 |
| | 30 | 2.994e-06 |

*,** and *** represents the significance level at 10,5,1 % levels respectively

The more robust test for autocorrelation was conducted, that is the Box Ljung. As a results from Table 4.5, the p-value corresponding to lag 10, 20, 30 and 36 are less than 0.01. Therefore the null hypothesis was rejected at 1 % level of significance and conclude that British Petroleum returns exhibit autocorrelation.

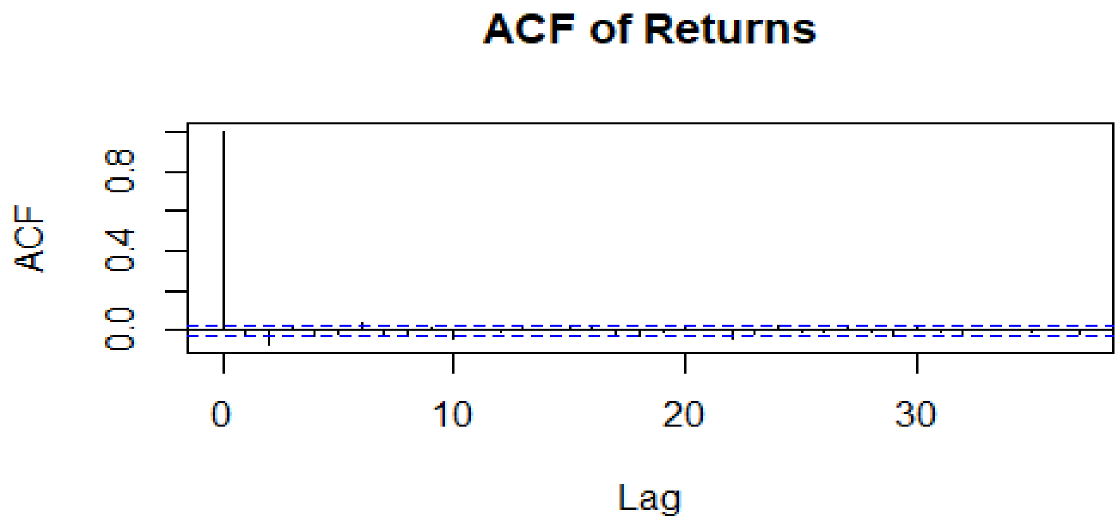



Figure 4.4: Autocovariance Function

## 4.8 Structural Break Tests

Chow test, SupF test and CUSUM are employed for checking the presence of structural breaks. This three tests have been widely used, it is advisable to used all for the consistence of results.

Table 4.6: Structural Breaks Tests of Returns

| Test | P value |
|---|---|
| Chow | 0.4817 |
| SupF | 0.9411 |
| OLS-based CUSUM | 0.9641 |

*,** and *** represents the significance level at 10,5,1 % levels respectively

The p-value for tests are greater than all levels of significance, hence we fail to reject the null hypothesis indicating that there no of structural breaks for British Petroleum returns.

## 4.9 Long Memory Tests

The auto-covariance function plot in Figure 4.3 of squared returns exhibit the long range dependence since spikes are decaying hyperbolic. The long memory test for British Petroleum squared returns are Periodogram estimator (GPH) and Whittle Estimator that are parametric and semi-parametric methods respectively. Table 4.7 shows Periodogram tests and whittle tests, that suggested the presence of long range dependence at 1% level of significance.

Table 4.7: Long Memory Tests of Returns

| Test | Hurst | P value |
|---|---|---|
| Periodogram Method (GPH) | | |
| Whittle Method | | |

*, ** and *** represents the significance level at 10, 5, 1 % levels respectively

## 4.10 Selection of Distributions of Innovation

One of the stylized facts of financial time series are heavy tailed distribution that are exhibited by returns. Therefore a vast number of distribution have been widely used to capture heavy tailed returns; namely normal (norm) distribution , skewed normal (snorm) distribution , student-t distribution (std), skewed student-t distribution (sstd), generalized error distribution (ged), skewed generalized error distribution (sged), normal inverse Gaussian (nig) distribution and Johnson's SU (jsu) distribution as distributions of innovations.

The most widely used criteria for selection are the Akaikes Information Criterion (AIC) and the Bayesian Information Criteria (BIC). It is advisable to use both information criteria to check the best distribution for heavy tailed returns since AIC overestimates the order of autoregression. Table 4.8 shows best distribution of inno-

Table 4.8: Heavy Tailed Distributions

| Model | IC | norm | snorm | std | sstd | ged | sged | nig | ghyp | jsu |
|---|---|---|---|---|---|---|---|---|---|---|
| GARCH | AIC | -5.502 | -5.502 | -5.542 | -5.542 | -5.539 | -5.539 | -5.542 | -5.542 | **-5.543** |
| | BIC | -5.497 | -5.496 | **-5.536** | -5.535 | -5.533 | -5.532 | -5.535 | -5.534 | -5.535 |
| IGARCH | AIC | -5.498 | -5.499 | **-5.541** | -5.541 | -5.537 | -5.537 | -5.541 | -5.541 | -5.541 |
| | BIC | -5.493 | -5.492 | **-5.535** | -5.534 | -5.531 | -5.530 | -5.533 | -5.532 | -5.534 |
| EGARCH | AIC | -5.513 | -5.515 | -5.551 | -5.552 | -5.548 | -5.548 | -5.552 | -5.552 | **-5.552** |
| | BIC | -5.507 | -5.507 | **-5.544** | -5.543 | -5.540 | -5.540 | -5.543 | -5.542 | -5.543 |
| TGARCH | AIC | -5.513 | -5.514 | -5.548 | -5.548 | -5.546 | -5.546 | -5.549 | -5.548 | **-5.549** |
| | BIC | -5.506 | -5.507 | **-5.541** | -5.540 | -5.538 | -5.537 | -5.540 | -5.538 | -5.540 |
| FIGARCH | AIC | -5.498 | -5.499 | **-5.541** | -5.541 | -5.537 | -5.537 | -5.541 | -5.541 | -5.541 |
| | BIC | -5.492 | -5.491 | **-5.534** | -5.533 | -5.529 | -5.528 | -5.532 | -5.531 | -5.532 |

vation for each GARCH-type model using AIC and BIC. The AIC favoured student-t distribution for IGARCH and FIGARCH models while prefered Johnson's SU distribution for EGARCH and TGARCH models. However, BIC suggested that student-t distribution as the best distribution of innovation for symmetric, asymmetric and long memory GARCH-type model. The student-t and Johnson's distribution were selected the best distribution of innovation because they have been standardized to ensure that their mean and standard deviation equal to zero and one respectively.

## 4.11 Parameters of GARCH Model

The British Petroleum returns are not constant over time, which suggests that returns are volatile. Therefore GARCH-type models were adopted so as to capture heteroskedasticity. The British Petroleum returns also exhibited heavy tailed returns, therefore eight distributions mentioned above were adopted to accommodate fat tails. The stylized facts such volatility clustering, leverage effects and Long range dependence are to be captured by symmetric GARCH-type models, asymmetric GARCH-type models and Long Memory GARCH-type models respectively. The study used standard GARCH model assuming Johnson's SU distribution and

IGARCH model assuming student-t distribution to accommodate volatility clustering. EGARCH model assuming Johnson's SU distribution and TGARCH model assuming Johnson's SU distribution were used to capture leverage effects.
The FIGARCH model assuming student-t distribution was adopted to capture the persistent long range dependence.

Table 4.9: Estimated Parameters of GARCH-type Model

| Model | GARCH | IGARCH | EGARCH | TGARCH | FIGARCH |
|---|---|---|---|---|---|
| Distribution | jsu | std | jsu | jsu | std |
| Constant($\mu$) | 0.000357 | 0.000462 | 0.000106 | 0.000188 | 0.000462 |
| Constant($\omega$) | 0.000002*** | 0.000001 | -0.084790*** | 0.000002 | 0.000001 *** |
| Constant ($\delta$) | | | | | 0.999999*** |
| ARCH Effect($\alpha$) | 0.056017*** | 0.059798*** | -0.055797*** | 0.019372*** | 0.019105 |
| GARCH Effect($\beta$) | 0.936407*** | 0.940202 | 0.989835 *** | 0.942913*** | 0.942887*** |
| Leverage Effect($\gamma$) | | | 0.102986 *** | 0.057906 *** | |
| Skewness | -0.094363 | | -0.147092** | -0.134521** | |
| Shape | 1.914104*** | 6.305334*** | 1.972886*** | 1.977069*** | 6.332448 |
| $\alpha + \beta$ | 0.992424 | 1 | 0.9340865 | 0.962285 | 0.961992 |
| ARCH Test | | | | | |
| LM (Lag 7) | 0.08188 | 0.1681 | 0.08919 | 0.1816 | 0.12669 |
| Ljung-Box (Lag 9) | 0.02861 | 0.09085 | 0.04038 | 0.2119 | 0.1187 |
| Information Criteria | | | | | |
| AIC | -5.5425 | -5.5419 | -5.5519 | -5.5486 | -5.5413 |
| Ranking | 1 | 2 | 1 | 2 | 1 |

*, ** and *** represents the significance level at 10, 5, 1 % levels respectively

### 4.11.1 Symmetric GARCH-type models

Table 4.9 shows two symmetric GARCH-type models that were fitted, that is GARCH model assuming Johnson's SU distribution and IGARCH model assuming student-t distribution. The AIC suggested that GARCH model assuming Johnson's SU distribution is the best symmetric GARCH-type model for British Petroleum returns. The sum of ARCH and GARCH term of the GARCH model from Table 4.9 is close to one indicating that volatility shocks are persistent. The p-value of Engle's LM test corresponding to GARCH model assuming Johnson's SU distribution is greater than 5%, therefore we fail to reject the null hypothesis and conclude that there is no heteroscedasticity at lag 7, which suggest that conditional volatility was absorbed by GARCH assuming Johnson's SU distribution. The second better model for symmetric GARCH-type models was IGARCH assuming student-t distribution. IGARCH model assuming student-t distribution has a sum of one for the ARCH and GARCH term shown in Table 4.9. This indicates that shocks due to volatility are highly persistent. The phenomena of highly persistent volatility, confirms volatility clustering in British Petroleum returns.

### 4.11.2 Asymmetric GARCH-type models

Table 4.9 shows two asymmetric GARCH-type models that were fitted namely; EGARCH model and TGARCH both assuming Johnson's SU distribution. AIC suggested that EGARCH model assuming Johnson's SU distribution was the best asymmetric GARCH-type model. The leverage effect ($\gamma$) is positive and statically significant which indicates that the presence of leverage effect and it reflects that bad news increases conditional volatility. The GARCH term for the EGARCH assuming Johnson's SU distribution has a value of 0.989835 which is close to one, indicating that volatility of British Petroleum returns took time to die. The sum of the ARCH and GARCH term is 0.934038 which is close to one suggesting that volatility shock are persistent. The TGARCH model under jsu shown in Table 9 indicates that bad or good news has a symmetric impact since $\gamma = 0$. The parameter $\gamma$ is significantly positive which suggests that leverage effect is present, which in turn implies that bad news increases volatility. The p-value of Engle's LM test for the EGARCH model assuming Johnson's SU distribution at lag 7, in Table 4.9 is greater than 5% level of significance. Therefore we fail to reject the null hypothesis that there is no heteroskedasticity and conclude that EGARCH model assuming Johnson's SU distribution was able to capture conditional volatility.

### 4.11.3 Long Memory GARCH-type models

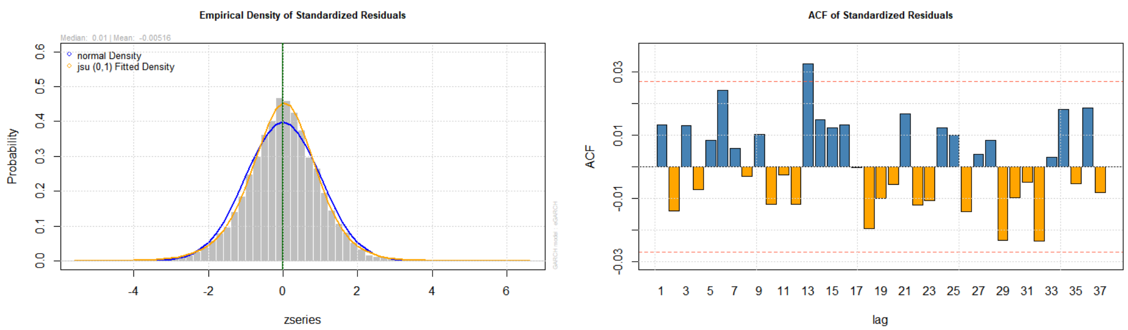
Table 4.9 shows FIGARCH model assuming student-t distribution is the only long memory GARCH-type model that was able to capture long range dependence for British Petroleum returns. Let $\delta$ denote long memory parameter in the volatility model. The long range dependence parameter is positive ($0 < \delta < 1$) and indicates a persistent long memory.The ARCH and GARCH term for FIGARCH assuming student-t distribution was 0.961992 indicating that effect of heteroskedasticity was captured.

## 4.12 Better Estimate of GARCH-type Model

Table 4.9 shows that EGARCH model assuming Johnson's SU distribution was able to capture stylized facts that were inherent in British Petroleum returns. The EGARCH assuming assuming Johnson's SU distribution is the best GARCH-type model that is able to capture volatility clustering and leverage effects, then followed by TGARCH model assuming Johnson's SU distribution. This is an indication that asymmetric GARCH-type model were superior in capturing the the financial stylized facts than symmetric GARCH-type model and long memeory GARCH-type of British Petroleum returns. The GARCH model assuming Johnson's SU distribution and IGARCH model assuming student-t distribution were the third and forth better model respectively to capture stylized facts of British Petroleum returns. This suggests that symmetric GARCH-type model was the second to capture better the stylized facts. The FIGARCH is not a good model in capturing the stylized facts of British Petroleum returns, which suggests that long memory was not able to capture the stylized facts.

### 4.12.1 Diagnostic Checking of the EGARCH Model

According to the AIC, the EGARCH model assuming Johnson's SU distribution in the errors perform better than all GARCH-type models considered in this study since it has the least AIC.



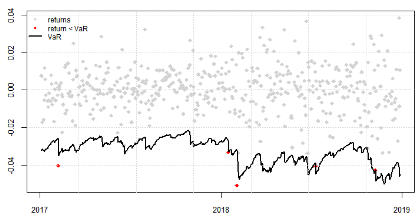

(a) Normality Test (b) ARCH Tests

Figure 4.5: Assessing Standardized Residuals of EGARCH

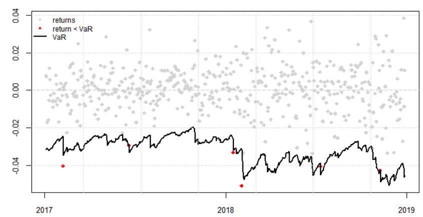
The density plot suggests the violation that of the fact that standardized residuals are normal distributed. On exploring the EGARCH model further we noted that heteroskedasticity was completely discarded reflected by Engle's LM test in Table 4.9. The ACF of standard residuals plot in Figure 4 exhibits no autocorrelation except for the one value that appear in lag 13, suggesting that Johnson's SU distribution was the best heavy tailed distribution.

## 4.13 Estimation Value at Risk

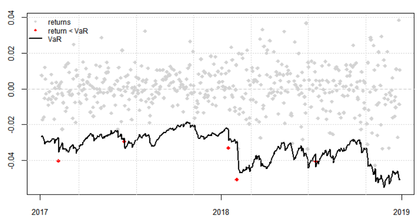
Figure 4.6 shows daily VaR calculations using the GARCH, IGARCH, EGARCH, TGARCH and FIGARCH model plotted against the British Petroleum returns. Figure 4.5 shows the downside 1% VaR in black line plotting against the returns with the red dots being the not captured volatility clustering , leverage effects and long range dependence.



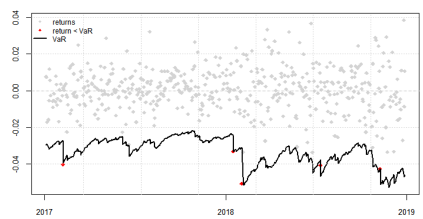

(a) GARCH-jsu



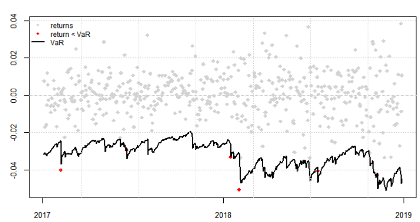

(b) IGARCH-std




(c) EGARCH-jsu




(d) TGARCH-jsu




(e) FIGARCH-std

Figure 4.6: VaR for GARCH-type models

The VaR calculated for GARCH-type models suggests a fairly stable trend that follows the bottom part of returns, however the TGARCH model seems to be able to capture volatility clustering and extreme events of British Petroleum returns. The GARCH, IGARCH, EGARCH and FIGARCH could not capture well extreme events especially in 2008/2009 when world was experiencing global crisis/ recession.

### 4.13.1 Back-testing Value at Risk (VaR)

To find out which type of GARCH-type of model best estimates volatility and VaR estimation of market risk. The Kupiec and the Christoffersen test performed using the VaR at 1% confidence level and the results summarized in Table 4.10. All the GARCH-type models passed the unconditional coverage and conditional test at 1% exceedance.

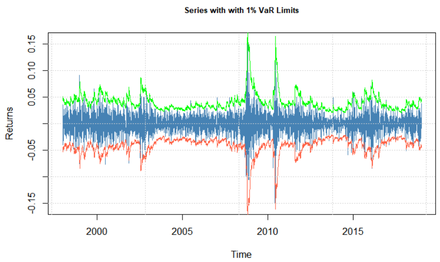
Table 4.10: VaR for GARCH Model

| Model | Distribution | Expected Exceed | Actual Exceed: | Christoffersen | Kupiec |
|---|---|---|---|---|---|
| GARCH | jsu | 5 | 5 | 0.951 | 1 |
| IGARCH | std | 5 | 6 | 0.845 | 0.663 |
| EGARCH | jsu | 5 | 5 | 0.951 | 0.397 |
| TGARCH | jsu | 5 | 5 | 0.951 | 1 |
| FIGARCH | std | 5 | 6 | 0.845 | 0.663 |

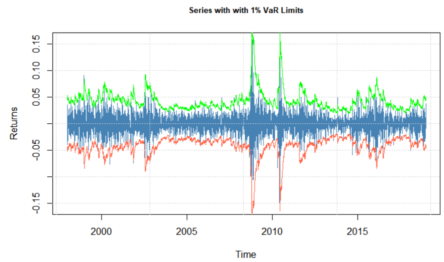
*,** and *** represents the significance level at 10, 5, 1 % levels respectively

The GARCH, EGARCH and TGARCH assuming Johnson's SU distribution had superior results than IGARCH and FIGARCH under student t-distribution in back-testing VaR.

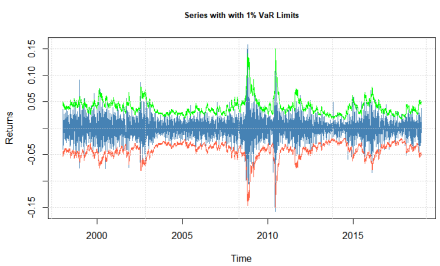
Figure 4.7 in blue line are British Petroleum returns and the red lines indicates VaR at 1%. Figure 4.7 also confirms which model was the best in terms of capturing volatility clustering, leverage effect and long memory.



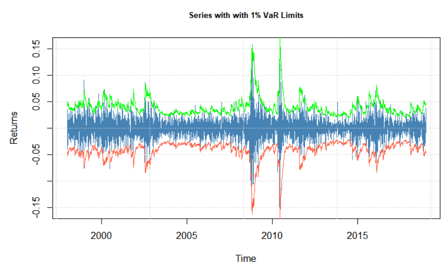

(a) GARCH-jsu



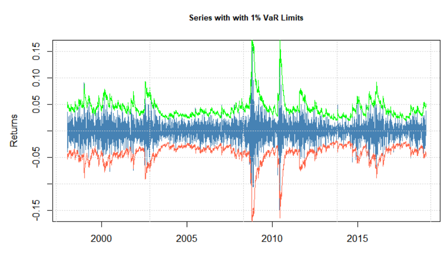

(b) IGARCH-std




(c) EGARCH-jsu




(d) TGARCH-jsu




(e) FIGARCH-std

Figure 4.7: Backtesting for GARCH-type models

The actual exceedance of GARCH, EGARCH and TGARCH is five, this implies that returns touches the VaR at 1% five times suggesting while IGARCH and FIGARCH the actual exceedance is 6, which implies that returns touches the VaR at 1% six times.

## 4.14 Forecast Evaluation

Forecast evaluation for GARCH-type models is a fundamental in determining forecasting accuracy of the models estimated. There are vast number of forecasts evaluation methods available in literature. This study employed three measure of accuracy, namely; MAE, MSE and RMSE. The best GARCH-type model must exhibit least prediction forecast error. The MAE, MSE and RMSE forecasting errors for the five GARCH-type model are reported in Table 4.11.

Table 4.11: Accuracy Measures of GARCH models

| Model | Distribution | MAE | MSE | RMSE |
|---|---|---|---|---|
| GARCH | jsu | 0.007980283 | 0.0001319092 | 0.01148517 |
| IGARCH | std | 0.007959403 | 0.0001311066 | 0.01145018 |
| EGARCH | jsu | 0.008030523 | 0.0001339297 | 0.0115728 |
| TGARCH | jsu | 0.008014143 | 0.0001332571 | 0.0115437 |
| FIGARCH | std | 0.007959383 | 0.0001311059 | 0.01145015 |

The FIGARCH assuming student-t as a distribution of innovation provides the most accurate forecasting power followed by IGARCH assuming student-t distribution. The third and fourth GARCH-type models with a better forecasting power are GARCH-type are GARCH assuming Johnson's SU distribution and TGARCH assuming Johnson's SU distribution respectively. Lastly the EGARCH assuming Johnson's SU distribution because it gives the least prediction error, this suggest that it a poor model in forecasting British Petroleum returns Results for accuracy measure for forecasting suggests that long memory GARCH-type model is the best in estimating the future volatility than symmetric GARCH-type models and asymmetric GARCH-type models. This gives an insight that FIGARCH was able to capture volatility clustering, leverage effects and long range dependence.

# Chapter 5

# Conclusions, Limitations and Recommendations

## 5.1 Conclusion

In this study, we analysed compound returns of British Petroleum Prices from 1998 to 2008. Our data analysis firstly checked for the presence of financial stylized facts. Time series plots of returns exhibited volatility clustering and leverage effects. Returns showed negative skewness and high kurtosis. This was confirmed by normality test that returns are not normal and exhibited heavy tailed distribution. The structural breaks tests showed that there was no breaks in the returns. Th auto-covariance plot of squared returns showed an exponentially decay which suggested the presence of long range dependence and long memory tests confirmed the presence of long range dependence.

The study focused on symmetric, asymmetric and long memory GARCH-type models to capture volatility clustering, leverage effects and long range dependence respectively. The heavy tailed distribution of returns were captured by eight distributions of innovations. The Akaike Information Criteria and Bayesian Information Criteria were used in selection of distributions of innovations that better explains the heavy tailed distribution. Three classes of GARCH-type models favoured student-t and Johnson's SU distribution which explains better the fat tailed returns. The best model fit for the symmetric GARCH-type model was the GARCH assuming Johnson's SU distribution, this suggested that it was able to capture volatility clustering and fat tails of returns. EGARCH assuming Johnson's SU distribution proved to be the best asymmetric GARCH-type model, which suggested that is was able to handle the three financial stylized facts that volatility clustering, leverage effect and heavy tailed returns. The long memory GARCH-type model that was able to accommodate long range dependence was the FIGARCH assuming student-t distribution. However, Akaike Information Criteria selected the EGARCH assuming Johnson's SU distribution as the best model that capture persistent stylized facts such volatility clustering and leverage effects, this suggested that long range dependence was not persistent on the returns. The model that showed stability on forecasting the Value at Risk for volatility clustering, leverage effects was the TGARCH assuming Johnson's SU distribution. This further suggests that long range dependence was not persistent as volatility clustering and leverage effects. However, Mean Absolute Error, Mean Square Error and Root Mean Square Error suggested that FIGARCH assuming student-t distribution as the best accurate forecasting GARCH-type model.

## 5.2 Limitation

Initially the study was focusing on three GARCH-type models for each of the following classes symmetric, asymmetric and long range dependence. The GARCH-M for symmetric GARCH-type models, Power GARCH for asymmetric GARCH-type modeland Fractionally Exponential GARCH and Hyperbolic GARCH for long memory

GARCH-type model failed to converge.

## 5.3 Recommendation

The GARCH-type model that was able to capture volatility clustering of British Petroleum returns as GARCH assuming Johnson's SU distribution. GARCH-type model that was able to handle the leverage effects of British Petroleum returns was EGARCH assuming Johnson's SU distribution. The long range dependence phenomena was captured by the FIGARCH assuming student-t distribution. Therefore, the recommendations for further studies is that they must consider Johnson's SU distribution in capturing heavy tailed returns since most researcher's have mostly considered student-t distribution. The investors who seek to minimise the risk associated to British Petroleum returns, TGARCH assuming Johnson's SU distribution is the best model for estimating stable Value at Risk. The FIGARCH assuming student-t distribution is suggested was better model in forecasting British Petroleum returns.

# Bibliography

[1] Alan Agresti. Categorical data analysis. john wiley & sons. *Inc., Publication*, 2002.

[2] Mary Louise Flint. *Pests of the garden and small farm: a growers guide to using less pesticide*, volume 3332. UCANR Publications, 2018.

[3] Wei Tang, Cliff I Davidson, Susan Finger, and Kirk Vance. Erosion of limestone building surfaces caused by wind-driven rain: 1. field measurements. *Atmospheric Environment*, 38(33):5589–5599, 2004.

[4] Cathy N Davidson. *Revolution and the word: The rise of the novel in America.* Oxford University Press, 2004.