UNIVERSITY OF GHANA

COLLEGE OF BASIC AND APPLIED SCIENCES

DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE

MODELLING AND FORECASTING EXCHANGE RATE IN GHANA: AN APPLICATION OF THE EXPONENTIAL GARCH AND BAYESIAN VECTOR AUTOREGRESSION MODELS

BY

JOSEPH WORLANYO NYADROH

(10477865)

THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA, LEGON IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF MPHIL STATISTICS DEGREE

JULY, 2023

DECLARATION

Candidate's Declaration

I hereby declare that this submission is my own research work for the award of the MPhil. Degree in Statistics. Furthermore, appropriate acknowledgement had been made where needed. It contains no content that has been previously published by another individual nor has been accepted for the award of any other university degree, to the best of my knowledge.

Signature..... Date.....26/10/2023.....

Joseph Worlanyo Nyadroh

(10477865)

Supervisors' Certification

We hereby certify that this thesis was prepared from the candidate's own research work and supervised in accordance with the guidelines of thesis laid down by the University of Ghana.

Signature..... Date.....26/10/2023.....

Dr. Ezekiel Nii Noye Nortey

(Supervisor)

Signature..... Date.....26/10/2023.....

Dr.Dennis Arku

(Co-Supervisor)

DEDICATION

This thesis is dedicated to my parents; siblings; and to the Department of Statistics and Actuarial Science.

ACKNOWLEDGEMENT

My profound gratitude goes first to Almighty God for seeing me through this journey and most especially my supervisors: Dr. Ezekiel Nii Noye Nortey and Dr. Dennis Arku for their efforts, patience and guidance throughout my research.

I am also grateful to Professor Lious Asiedu (current Head Of Department), Professor Kwabena Doku-Amponsah (Immediate past Head of Department) and all the lecturers, staffs and students at the Department of Statistics and Actuarial Science for facilitating my academic work here in the university.

I appreciate my parents: Mr Helmut Kwasi Nyadroh and Mrs. Lucy Bockor Nyadroh; and my siblings: Clement, Eselasi and Jacob for their immense support throughout my studies.

And also, to my boss Dr. Felix O. Mettle for his guidance and assistance. Not forgetting my dear mates especially Mr. Emmanuel Annor, Mohammed Kamil, Ama Frimpomaa Boateng, Stephen Aboagye and Seth Anfo for their support and words of encouragement.

ABSTRACT

Ghana had faced a macroeconomic problem of exchange rate for a long period of time and this problem had somehow slowed the economic growth. The exchange rate has been one of the major economic challenges facing most countries in the world especially African countries including Ghana. Therefore, forecasting and modelling the nature of the exchange rate between the Cedi and the US dollar in Ghana is very important in order for the government to design economic strategies and effective monetary policies to combat any unexpected hike in the exchange rate. This research studied the variances in the exchange rate data in Ghana and found statistical models to represent and forecast the variance associated with it. Using monthly data on the exchange rates from January, 2008 to March, 2022, the Bayesian Vector Autoregression (BVAR) model gave good forecast of the rates with a BIC of -162.491 and an RMSE of 0.017. Based on the Bayesian VAR model, we also forecasted the rates outside the sample period (January, 2019 to December, 2021). The forecasted results showed consistency in the actual data. The study concluded that multivariate models such as the Bayesian VAR was good in forecasting the rates and was also able to assess the impact of other variables on the exchange rate data. The study recommends the use of the Bayesian VAR model in analyzing macroeconomic data. Also, the use of a complex model such as the GJR-GARCH could go a step further in dealing with the variances within the residuals of the exchange rate data, while assuming a prior distribution that fits the data for the parameters of the Bayesian VAR could give better forecasts of the exchange rate.

Contents

List of Tables

List of Figures

LIST OF ABBREVIATION

ARCH ..... Autoregressive Conditional Heteroscedasticity

ADF ..... Augmented Dickey-Fuller

AR ..... Autoregressive

ACF ..... Autocorrelation Function

AIC ..... Akaike Information Criterion

ARMA ..... Autoregressive Moving Average

ARIMA ..... Autoregressive Integrated Moving Average

BIC ..... Bayesian Information Criteria

BVAR ..... Bayesian Vector Autoregression

CPI ..... Consumer Price Index

EGARCH ..... Exponential Generalized Autoregressive Conditional Heteroscedasticity

ERP ..... Economic Recovery Program

EXRATE ..... Exchange rate

FEVD ..... Forecast Error Variance Decomposition

GARCH ..... General Autoregressive Conditional Heteroscedasticity

GDP ..... Gross Domestic Product

H-Q ..... Hannan-Quinn

HIPC ..... Highly Indebted Poorest Country

IID ..... Independent and Identically Distributed

INF ..... Inflation

IRF ..... Impulse Response Function

KPSS ..... Kwiatkowski-Phillips-Schmidt-Shin

LM ..... Langrange Multiplier

MLE ..... Maximum Likelihood Estimation

MA ..... Moving Average

MSE ..... Mean Square Error

MAE ..... Mean Absolute Error

MAPE ..... Mean Absolute Percentage Error

MCMC ..... Markov Chain Monte Carlo

OLS ..... Ordinary Least Square

PACF ..... Partial Autocorrelation Function

Q-Q plot ..... Quantile-Quantile plot

RMSE ..... Root Mean Square Error

SSR_o ..... Regression Sum of Square

SSR_1 ..... Residual Sum of Square

SAP ..... Structural Adjustment Program

TGARCH ..... Threshold Genralized Autoregressive Conditional Heteroscedasticity

US ..... United States

VAR ..... Vector Autoregression

WACB ..... West African Currency Board

Chapter 1

INTRODUCTION

1.1 Background of the Study

During the days of independence in 1957, Ghana was doing well than many of its neighbouring countries in the West African region. Its cocoa sector was booming and it was the major supplier of cocoa to the world with substantial amount of gold. It had enough foreign exchange reserves. Its labour force were skilled with some improved infrastructure. Ghana achieved a middle income status with good per capita income (Aryeetey & Fenny, 2017).

Ghana took off as a country well endowed with natural resources with hopes that situations would become even more better in the years after Independence. This was because its first leader Dr. Kwame Nkrumah was very moved to redeem his citizens from poverty. But it turned otherwise and 25 years after independence, the country's economy had ended up in shambles due to a weakening trade sector. The nation then went through some robust reforms which brought about appreciable growth in the economy. Upon all these reforms, the exchange rate policy has still been a challenge and this has taken a toll on the external performance of the country (Laryea & Senadza, 2017).

For a nation's market to flourish, it ought to have an exchange rate that is competitive. This is the main mechanism for manipulating the expenditure trends of a nation. It had become a prominent feature in the Ghanaian economy and all efforts to control it became the focus of the whole transformation process. Policy reforms on the rate of exchange have not been in line with the plans put

forward in the beginning.

The government in its efforts tried possible means to acquire enough foreign exchange and make exports rise. In its quest, the government planned to devaluate the cedi, adopt a market determined rate of exchange, license forex bereaus to operate as alternate markets in order to preserve a flexible exchange rate system. These reforms of the government actually helped to attain some improvements in exports, but it was still heavily dependent on import (Dordunoo, 1994) as cited in (Laryea & Senadza, 2017).

When the Bretton Woods agreement was introduced in 1973, it made countries to implement the floating system of exchange rate which led currencies of different nations to either appreciate or depreciate. A continuous depreciation of the currencies of many countries as a result of the floating system of exchange rate brought about some level of awareness.

Over the past few years, Ghana's currency has been consistently depreciating. Nigeria and Kenya have also experienced dynamics of their currency for some time. This has not been the situation for Ghana. In the year 2007, the Central Bank redenominated the cedi which led to the cedi being at par with the US dollar. This effort by the Central Bank was not sustainable. Presently, a dollar equals 11 cedis .

This prolonged depreciation of the cedi has caught the attention of policy makers. Since the international market plays an important part in the economy, by supporting macroeconomic performance and economic growth. Ghana is heavily dependent on import and a continuous fall in the value of its currency would lead to a rise in the general prices of goods and services, a decrease in investment and balance of payment issues (Agyemeng, 2021).

The exchange rate is characterized by either an appreciation, depreciation or stability. An appreciation of the exchange rate leads to overvaluation and makes imports to be less expensive whilst exports continues to be expensive, thereby making the international competitiveness of the country to reduce (Takaendesa, 2006). A depreciation of the currency also leads to difficult terms of trade, makes imports to be expensive whilst exports become cheaper. Chowdhurry (1999) made it clear that a stable currency fosters economic growth and is perceived desirable. It is on this basis that policy makers develop programmes aimed at stabilizing the cedi .

Mordi (2006) claimed that the dynamics of the exchange rate affects inflation, export competitiveness, international confidence and balance of payment equilibrium. This study will contribute to the Central Bank's efforts to combat the depreciation of the cedi. It is only when we are able to detect in advance what the exchange rate movement would be that the Central Bank can come up with ways to stop its depreciation.

A lot of researchers have tried to develop statistical models that may help determine and describe the nature of these macroeconomic variables. For instance, a two-stage least square econometric method was utilised to research the relation between interest rate and exchange rate by Tomiwa (2018) in Japan from (Antwi et al., 2020). He considered the supply of money and GDP as the main macroeconomic indicators of the relation between interest rate and exchange rate. It gave evidence of a good relation in Japan's trade and currency.

Also a simple monetary model for obtaining exchange rate and a co-integration analysis was adopted to examine the main factors that come into play in the cedi to dollar rates (Mumuni & Owusu-Afriyie, 2004). The simple model was

coupled with political factors to bring out any potential effect on the rate of exchange.

Furthermore, co-integration analysis method was employed by Owusu-Antwi et al. (2016) to detect the factors impacting the exchange rate behaviour in Ghana. It scrutinized the exact indicators that determined the exchange rate. It highlighted two main indicators which are government expenditure and historical exchange rates. However, in a longer period of time, Consumer Price Index (CPI), nominal GDP, domestic credit, imports and government expenditure come into play.

From Doan et al. (1984), the Bayesian procedure to the estimation of vector autoregression (VARs) was used. Using VARs in empirical study of macroeconomic variables brought about some controversies. Most forecasting models are traditionally formulated as simultaneous equations in structural form. This makes them not too suitable for forecasting. VARs gives us another option to structural macroeconomic models for forecasting purposes (Kenny et al., 1998). However, VARs have also been downplayed as they have inadequate theoretical backing over the use of theory as a check in choosing which variables to involve in the analysis .

Doan et al. (1984) again made an effort to make the forecasting performance of unrestricted VARs better by proposing that they could be estimated using Bayesian methods which accommodates prior information which may be available to the researcher. There is adequate empirical prove which gives us the idea that Bayesian VARs generates forecasts which are far better when compared to univariate time series models or a more sophisticated model. It is from this Bayesian method of estimating parameters in VARs that is being used in tbis study. Hence, it requires some investigation of the situation in the Ghanaian

economic setting.

The end results of this study serve as addition to exiting literature and bring renewed ideas on the exchange rate and other related macroeconomic indicators. This would help the Central Bank to monitor the impact these macroeconomic variables have on the exchange rate and help forecast the rate in an improved way. This study would also help stakeholders get in-depth knowledge on exchange rate for planning.

1.2 Statement of Problem

Ghana's economy is influenced by the exchange rate and its volatility. In both economic and statistical literature, fluctuations in the exchange market have received a lot of attention. The rise and fall of exchange rates over time is referred to as exchange rate volatility.

Ghana's economy has been vulnerable to fluctuations in the cedi to dollar rate and its forecasting. Statisticians use a variety of statistical approaches for forecasting and assessing model performance based on statistical power in order to arrive at a solid forecast.

To anticipate the exchange rate, we will need a statistical model with a small margin of error coupled with a reliable statistical test and this has been the subject of several research.

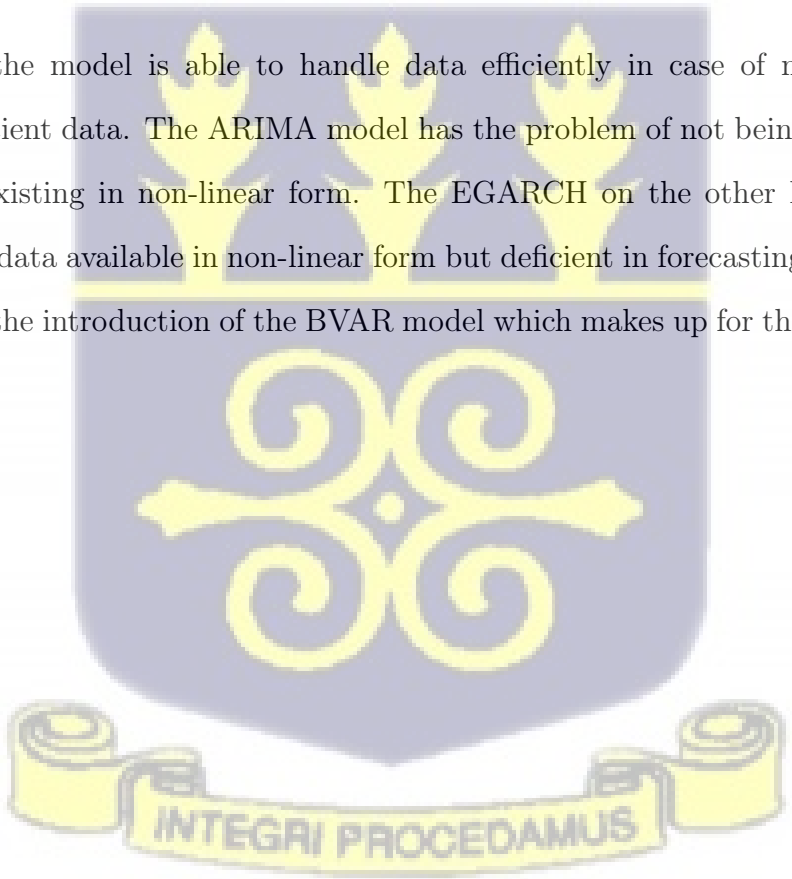
Several studies have been conducted on the ARIMA, Exponential GARCH and Vector Autoregression (VAR) models been used in forecasting exchange rates and other macroeconomic variables in Ghana, but not many studies have been conducted on the usage of the Bayesian Vector Autoregression (BVAR) model in forecasting macroeconomic variables such as the exchange rate, inflation and

GDP in Ghana.

The several parameters of the VAR model have become one of the model's key flaws, despite the fact that it is succinct and unambiguous and frequently employed in macroeconomic projections. The VAR model's over-parameterization flaw is solved by the Bayesian Vector Autoregression (BVAR) model.

To avoid the over-parameterization problem, the BVAR model leverages statistical features of variables as prior knowledge for the VAR model parameters. It has a very low forecast error with good forecast capability.

Also, the model is able to handle data efficiently in case of missing data or insufficient data. The ARIMA model has the problem of not being able to model data existing in non-linear form. The EGARCH on the other hand is able to model data available in non-linear form but deficient in forecasting appropriately, hence the introduction of the BVAR model which makes up for those deficiencies.



1.3 Objectives of the Study

This study seeks to determine an appropriate model to forecast the exchange rate in Ghana. The following specific objectives would be considered in this study:

1. Derive the best forecasting model for the conditional variance of the exchange rate.
2. Use a Bayesian VAR to fit and predict the exchange rate data.
3. Examine the relationship between the exchange rate and other economic variables.

1.4 Significance of the Study

The results from this study would be vital to stakeholders and policy makers such as government agencies, businesses, the entire public as well as academia and researchers because of the following reasons;

Foremost, by identifying an appropriate model a robust forecast of the values of economic variables such as the exchange rate, inflation etc. will be essential in the planning and decision-making activities of the entities named above.

Secondly, the outcome from this study will be beneficial to researchers and academia by adding to existing literature in the case of developing countries. The discovery of the model under investigation will lead to an expansion of research in this space.

1.5 Limitations of the Study

The foundation of every financial analysis is obtaining data on a series of past records in the particular subject area of study. This is a requirement for identifying the kind of problem in the area of study and discovering measures to address those problems.

In Ghana, accurate financial data is usually hard to come by and even if it is acquired, the information is not adequate enough due to the fact that not all the events were recorded in the databases of the relevant institutions.

Also, a mismatch of records obtained from the various agencies involved in collecting such data would be worrying. Despite all these, it is wise to concede that the data provided by these agencies might be under-recorded.

However, there is sufficient evidence from researchers who have used data from the Central Bank of the reliability and dependability of the data which contains salient features needed for similar studies.

Furthermore, getting a suitable prior distribution for the parameters of the Bayesian VAR model was a bit challenging, but the prior distribution assumed for the parameters of the model was appropriate for the data at hand, and had been tested in several studies by other researchers.

The EGARCH on the other hand was chosen among other complex models to account for the volatilities in this study since it had the ability to capture the assymmetries associated with the volatile nature of the rates. There might be other complex models that could perform better but the EGARCH had been tested in several studies to be good, and is specifically selected as suitable and

adequate for this particular study.

1.6 Organization of Study

The study is organized into five chapters. Chapter one contains the background, problem statement, the study objectives, significance of the study and the organization of the study.

Chapter two presents the introduction, history behind the exchange rate, some empirical studies on exchange rate in Ghana, review of some models used in studying time series, studies involving BVAR and other models, and a conclusion on the literature.

Chapter three lists the introduction, the research design used, the source of data for the study, some stationarity tests used, a detailed description of the Bayesian Vector Autoregression (BVAR) model, the impulse response function, forecast error variance decomposition, the MCMC method, some volatility models and finally the diagnostic and evaluation methods used on the model.

The chapter four looks at the analytical methods conducted on the data as discussed in chapter three. As well as relevant statistical tests performed on the data. Finally chapter five presents the research summary, conclusion and recommendations.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

This chapter generally looks at what had been done in the past by other researchers in relation to this particular study. It critically looks at the empirical approach to dealing with situations, and how the whole exchange rate administration started in Ghana. Then a final declaration of what this study intends to bring on board.

2.2 History of the Exchange Rate in Ghana

Exchange rate as defined in the Cambridge dictionary is the value a person receives for changing the money of one country with that of another country. It is the value one receives in a foreign currency of equal amount in his local currency (Alagidede et al., 2013). It is understood by many that the exchange rate is a volatile indicator in the economy among all other economic variables.

Stakeholders have raised lots of concern about how to manage the exchange rate (Senadza & Diaba, 2017). They observed from their study that as the gold market price fell, it became an issue managing the exchange rate. Concerns on the management of the exchange rate began early 1930 from the huge loss on the gold market price, which led to the development of the Bretton Wood agreement leading to the adjustment of currencies in 1940 .

Rapid variability in the value of a nation's currency impacts businesses as investors in the economy would prefer economic stability. When the exchange rate remains stable in an economy, it brings about economic growth. Many nations adopt one system or the other in managing their rate of exchange. Countries in early times adopted the flexible exchange rate system until 1970 where a lot of countries began using the fixed system.

After independence, the fixed system of exchange was still used by Ghana from the arrangement under the colonial agreement. To monitor how the currency was running in the system, the British in 1912 formed a board for their West African colonies to use the West African Pound and a fixed rate was determined against the sterling.

Ghana could not follow the stringent rules of the West African Currency Board (WACB), hence opted out and adopted the cedi in 1963 (Jebuni et al., 1991). The early stages of the adoption of the cedi saw the cedi being almost at par with the dollar and the nation continued with the use of the fixed exchange system.

When the Structural Adjustment Program was introduced, Ghana shifted to the floating exchange system where the rates were determined by the market. Forex bureaus were introduced in 1988 to serve as a mediator between the legal and illegal markets which led to the inclusion of the black market. With all things being equal, the floating exchange rate system led to increase demand for foreign currency which resulted in the depreciation of the cedi (Odutayo, 2015).

In 1990, the Government approved 180 forex bureaus to operate (Dordunoo, 1994; Bhasin, 2004). Then in 1992, the two-window exchange systems that is the fixed and floating were merged, and the market began to operate an interbank

whole sale system (Jebuni & Accra, 2006). With only few active firms in the exchange market, it was not enough. The Central Bank has been the most authoritative agency in the market. There are four divisions of the market. These are:

- The interbank market where banks trade among themselves
- Forex bureau which serves individuals, tourists, SMEs, etc.
- The corporate market through which transactions between banks and their customers are conducted and
- The unofficial market which comprises of corporations that price their products and services at their own-determined exchange rate

In recent times, countries have used variety of exchange rate systems, some are the gold standard (fixed exchange rate), the Bretton Woods system, and the 'managed' floating exchange system (Argy, 2013).

2.3 Empirical Studies on the Exchange Rate

Due to its economic and political consequences on a nation, the exchange rate is among key economic adjustment instruments as well as one of the most challenging economic policy tools (Jebuni & Accra, 2006). Mussa (1986) found that, the variables that influence exchange rates in an economy are mostly determined by the country's exchange rate system, either flexible or fixed.

Many economists are confused by the behaviour of the exchange rate in a floating system (Stockman, 1987). He gave detailed account of the behavior of the exchange rate and how they are set. According to him, there is demand on foreign exchange because they are needed to purchase goods and services. He also said that exchange rate projections should be made in anticipation of changes in terms of trade or factors related to trade, as well as speculations in

the inflationary rate.

This claim was backed by research done by (Mumuni & Owusu-Afriyie, 2004). They used the monetary method to examine the setting of the cedi to dollar rate in Ghana. They noticed that the supply of domestic and foreign money also accounts for the dynamism in the exchange rate. They declared in their work that the rate is mainly set by other variables that impact exchange rate and peoples speculations from past trends of the rate.

Stakeholders are much concerned about developments in exchange rates (Bartram et al., 2005). They suggest attention be given to the rates as it is an important element in the administration of the economy. Countries that adopted the floating system of exchange are more likely to be affected by any slight change in the rates, which would in turn impact their balance of payments.

2.4 Review of some models used in Time Series Analysis

In the mid 1920's, Gottman and Ringland (1981) developed a stochastic way of handling time series. An Autoregressive (AR) model came into limelight when Yule (1927) experimented on the Wolfer's sunspot data. Within that same period Slutsky (1927) from Nerlove and Diebold (1990) discovered the Moving Average (MA) model when experimenting on white noise series. In the 1970s, Box and Jenkins established the Autoregressive Moving Average (ARMA) model and later extended it to the Autoregressive Integrated Moving Average (ARIMA) model.

In 1943, Mann and Wald gave a theoretical layout on the approximation of Autoregressive AR (p) parameters using the least squares procedure. Quenouille

(1947) from Zellner et al. (1974) developed a test for parameters of the Autoregressive AR (p) models and further extended it to Moving Average (MA) models.

A method for estimating the order of the Autoregressive model (AR) model together with its parameters was due to the contribution of (Anderson, 1978). Later Box and Jenkins provided an extension which led to an estimated likelihood solution for ARMA (p,d) models. But the accurate version of the likelihood method for calculating the parameters of Moving Average MA (q) models and that of the ARMA (p,d) models were that of the efforts of Newbold (1974) from (Jones, 1980).

The selection of ARIMA model based on an information criterion was first conducted by Aikaike in 1974. A model with least AIC gives minimal mean square error with good forecast. Schwartz (1978) argued that using AIC's does not give consistent results in terms of probability hence proposed a Bayesian Information Criterion (BIC) for the selection from (Liddle, 2007).

An accurate measure of the parameters of an ARIMA model in state-space form was established with the help of (Harvey & Phillips, 1982). The state-space form are also referred to as Structural Time Series (STS) models. Durbin and Koopman (2001) supported the development of this method and claimed that it has benefits over the ARIMA as cited in (Commandeur et al., 2011).

2.5 Application of Volatility Models in some Studies

Volatility as defined in the Cambridge dictionary is the likelihood of something changing suddenly or unexpectedly. In this particular study, there are volatilities

in the exchange rate data, so a mathematical model would be developed to best describe the data and forecast future rates of exchange.

Laryea et al. (2016) defined volatility as an instability or uncertainty. He took volatility as a determinant of risk, no matter the phenomenon being discussed which serves as a source of addition to economic decisions. Variations in the exchange rate reflects in traded goods worldwide .

When the use of flexible exchange rate system gained prominence in 1973, a lot of studies have since been conducted using volatility models in recent times. The ARCH model had been a key tool in studying the spikes in data.

Engel (1982) wanted to model a time-varying conditional variance and came up with the ARCH model in (Kim & Nelson, 1989). In his discovery, he used the maximum likelihood estimation procedure under normality assumption. Later on it was developed into Generalized ARCH by (Bollerslev, 1986).

Mark (1988) used the Generalized Method of Moments to estimate the ARCH model. Robinson (1987) together with Gallant and Nychka (1987) in their estimation of ARCH model employed a non-parametric and semi-parametric approach from (Kozumi & Polasek, 2000; Takada, 2008).

In the fitting of a model to the fluctuations in the Istanbul gold exchange, Gencer and Musolglu (2014) in Aviah (2018) tried lots of GARCH models which accounted for longer intervals in the conditional variance and assymmetry. A comparison of the EGARCH and other GARCH models were made and the study finalized that the models that accomodated assymmetry and long dependence were efficient in prediction.

A study on the Malaysian stock market was conducted to find out if the symmetric and asymmetric GARCH were good (Lim & Sek, 2013). The models were subjected to seasons, that is, before, during and after tribulations. It concluded that the symmetric models were the best before and after the tribulations, whilst the asymmetric model did best during the tribulations with higher spikes.

EGARCH (1,2) turned out to be the best when it was used in modelling inflation rates in Ghana using monthly data (Nortey et al., 2014). ARIMA $(3,1,2)(0,0,0)_{12}$ was chosen as the best ARIMA model which was modelled together with the volatility models.

In the forecasting of the volatilities in the exchange rate between the dollar and Naira, the TGARCH model proved best when the TS-GARCH, TGARCH, GJR-GARCH and GARCH models were employed in the study (Iyoboyi & Muftau, 2014) .

2.6 Studies involving BVAR and other Time Series Models

The BVAR had been vastly used in advanced countries in the world but its usage in forecasting economic variables in a developing country like Ghana had seen no use. Lots of research done elsewhere adopted this method like the one done by (Litterman, 1986). The BVAR was applied to project real GDP, unemployment and inflation.

Litterman (1986) in his experiment made an effort to curtail the over-parameterization problem which led him to make a flexible assumption on the

coefficients. He assumed a normal distribution for the prior parameter. He also used the Theil's mixed estimation method as in Doan (2007) to approximate the parameters according to (Krol, 2010).

The variations in the Romanian economy was studied when Caraiani (2010) applied the BVAR to project its outcome. He obtained different forms of the BVAR and compared its forecasting abilities with that of the OLS and unrestricted VAR. He assumed the Litterman's prior which considers the parameters as following a random walk. He then concluded that the BVAR is the best in terms of forecasting .

Again, Spulbar and Birau (2023) in their quest to investigate shocks in the Romanian economy, relied on the BVAR model in their research. They also wanted to identify the origin of the shocks in the economy over the past 10 years. The Minnesota prior was taken into consideration for the study.

The BVAR emerged as the best model when Ramos (1996) predicted the market share of a leading car company in Portugal. In his work, he compared the foreacasting abilities of VAR, ARIMA and BVAR models. His choice of prior was contingent on the accuracy of the out-of-sample forecasts. He further made it clear that the model possessed some vital information such as impulse response function and decomposition of variance for making projections.

The tax revenue of California was forecasted using BVAR when Krol (2010) applied this technique in his study. He wanted to know if the BVAR would be superior to the VAR and other univariate models. After experimentation, he observed that the RMSE of the BVAR was less compared to the VAR model (Krol, 2010).

The Litterman prior was used by Yao (2011) in his work. He obtained lags of the economic variables and assumed they followed a random walk as well.

The Bayesian VAR turned out to perform excellently in the investigation done by (Carriero et al., 2012). They wanted to determine if BVAR could do better when a constant and drifting coefficients were considered in projecting key fiscal variables. They tried data from UK, US, Germany and France. The result they obtained proved that multivariate models do well in projection than their univariate counterpart.

There is not adequate literature on VAR using unequal time series but relevant background information can be found on Granger and Newbold (1986) from (Marcellino et.al., 2003; Banerjee et.al., 2005; Tutberidze & Japaridze, 2021; Stock & Watson, 2017) .

2.7 Theoretical Framework

2.7.1 Components of a Time Series

Components of time series are the factors that bring about changes in a time series. An important procedure in selecting the right modeling and forecasting procedure is to consider the type of data patterns shown from the time series graphs of the time plots. The sources of variation in terms of patterns in time series data are mostly classified into four main components:

- Trend
- Seasonal
- Cyclic
- Irregualr fluctuations

2.7.2 Autocorrelation function

Autocorrelation refers to the correlation of a time series with its own past and future values. For stationary processes, autocorrelation between any two observations only depends on the time lag h between them. Let us consider a stationary process of y_t and y_{t-h}.Now, the correlation coefficient between y_t and y_{t-h} is called the lag autocorrelation of y_t.

The autocorrelation of y_t denoted by ρ is generally given by

$$\rho = Corr(y_t, y_{t-h}) = \frac{cov(y_t, y_{t-h})}{\sqrt{var(y_t)var(y_{t-h})}} = \frac{Cov(y_t, y_{t-h})}{Var(y_t)} = \frac{Cov(y_t, y_{t-h})}{\gamma_o}$$

Let's define $Cov(y_t, y_{t-h}) = \gamma_h$

Therefore, $\rho = \frac{\gamma_h}{\gamma_o}$, the denominator γ_o is the lag 0 covariance, that is the unconditional variance of the process. The autocorrelation (ρ_h) lies betweeen -1 and 1 $(-1 \leq \rho_h \leq 1)$ and for any stationary process, $\rho_o \geq |\rho_h|$ for any h.

2.7.3 Partial Autocorrelation Function

Partial autocorrelation function measures the degree of association between y_t and y_{t-h} when the effect of other time lags on Y are held constant. It is basically the autocorrelation between y_t and y_{t-h} after removing any linear dependence on $y_1, y_2, y_3, ..., y_{t-h}, y_{th+1}$. The partial autocorrelation is denoted by $\psi_h h$. This is given by:

$$\psi_h h = Corr[y_t - E(y_t/y_{t-1}, ..., y_{t-k+1}), y_{t-k}]$$

where $E(y_{t-1}, ..., y_{t-k+1}$ is the minimum mean squared error predictor of y_t by $y_{t-1}, ..., y_{t-k+1}$.

2.7.4 Models of Time Series

2.7.4.1 Autoregressive Model (AR)

An AR (p) (Autoregressive of order P) model is a discrete time linear equation with noise, of the form:

$$X_t = \alpha_1 X_1 + \alpha_2 X_2 + ... + \alpha_p X_p + \epsilon_t$$

Here p is the order, $\alpha_1, \alpha_2, \alpha_3, ..., \alpha_p$ are the parameters, ϵ_t is an error term, usually a white noise with intensity ς^2. The model is considered either on integers $t \in Z$, thus without initial conditions, or on the non-negative integers $t \in N$. In this case, the relation above starts from $t = p$ and some initial condition $X_0, X_1, X_2, X_3, ..., X_{p-1}$ must be specified.

2.7.4.2 Moving Average Models (MA)

A model in which future values are forecast based on linear combination of past forecast errors is called moving average model.

A moving average model of order q, with mean zero, denoted by MA (q) is generally given as

$$x_t = \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + ... + \theta_p \epsilon_{t-p} + \epsilon_t$$

$$x_t = (\theta_1 L + \theta_2 L^2 + \theta_3 L^3 + ... + \theta_p L^p)\epsilon_t$$

$$x_t = \theta(L)\epsilon_t \ ,$$

where $\theta_1, \theta_2, \theta_3, ..., \theta_p$ are the moving average model parameters and ϵ_t is assumed to be Gaussian white noise.

In this general case, the PACF damps down and the ACF cuts off after q lags. An MA (q) model is neccesarily stationary if q is finite.

As MA (q) is said to be invertible if $\theta(L) = 0$ all lie outside the unit circle. A finite AR is always invertible.

2.7.4.3 Autoregressive Moving Average (ARMA) Model

Autoregressive moving average process involves the combination of the autoregressive and moving avearge processes of order p and q respectively. The time X_t is an ARMA (p,q) model if it is stationary and also if X_t is given by

$$X_t = \psi_1 x_{t-1} + \psi_2 x_{t-2} + ... + \psi_p x_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \theta_3 \epsilon_{t-3} + ... + \theta_q \epsilon_{t-q}$$

$$X_t = \psi_1 x_{t-1} + \theta(L)\epsilon_t$$

$$\psi(L)X_t = \theta(L)\epsilon_t$$

An ARMA process is stationary if the roots of $\psi(L)$ all lies outside the unit circle and invertible if the roots of $\theta(L)$ all lie outside the unit circle.

2.7.4.4 Seasonal Autoregressive Integrated Moving Average (SARIMA) Models

This type of models are essential in fitting or describing series with dynamic mean and some other statistic for a specified season across the years. It is developed from the usual ARMA and ARIMA models.

Say, a series has monthly observations and we would like to study some relation among a particular month in the series.We need to have an exact statistical representation of the components in the model.Our model can take the form;

$$\phi(B^s)\Delta_s^D z_t^\lambda = \theta(B^s)\alpha_t$$

where $\phi(B^s)$ and $\theta(B^s)$ represents seasonal AR and MA functions respectively, and α_t is a residual series which may contain non-seasonal correlation.

The AR and MA functions are specified to describe correlation existing between matched seasons.Seasonal AR functions is specified as;

$$\phi(B^s) = 1 - \phi_1 B^s - \phi_2 B^{2s} - ... - \phi_p B^{ps}$$

where ϕ_i shows the position of the AR coefficient and p is the rank of the AR function. Since the power of individual differencing function is often an integer multiple of s,only outcomes within individual seasons are linked to one another when this function is used.

To narrate the connection between the residuals within a given period. The seasonal MA function is specified using

$$\theta(B^s) = 1 - \theta_1 B^s - \theta_2^{2s} - ... - \theta_q^{QS}$$

where θ_i is the position of the MA coeffient and Q is the rank of the MA function. Since the power of B in $\theta(B^s)$ are often integer multiples of s, the residuals in matched seasons are connected using $\theta(B^s)$.

2.7.5 Time Series Volatility Models

2.7.6 ARCH (p) Model

The distinction of the ARCH model from the earlier time series model is the idea that taking the variance of the nuisance component at any particular time t relies on the squared nuisance component of previous periods of time.

A model invented by Engle (1990) is essential and competent in simultaneously using the mean and variance of a time series when the conditional variance fluctuates. It is helpful in financial and econometric modelling of conditional heteroscedasticity and volatility clustering.

Let the variance of a_t be conditioned on a_{t-i} which is the mean corrected return and $\epsilon_t \sim WN(0,1)$ and $H_t = a_1, a_2, ..., a_{t-1}$ be data on a_t at any particular time t.Then the ARCH (p) model is

$$x_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \beta_0 + \beta_1 x_{t-1}^2 + \beta_2 x_{t-2}^2 + ... + \beta_m x_{t-m}^2$$

where $\beta_0 \neq 0, \beta_i \geq 0, i = 1, ..., p$ and

$$E(a_t/H_t) = E[E(a_t/H_t)] = E[\sigma_t E(\epsilon_t)] = 0$$

$$V(a_t/H_t) = E(a_t^2) = \sigma_t^2 = \beta_0 + \sum \beta_i a_{t-i}^2$$

and the disturbance term ϵ_t is such that $E(a_t/H_t) = 0$ and $V(a_t/H_t) = 1$

2.7.7 Testing for ARCH effects

When conditional heteroscedasticity exits, there is an ARCH effect present. A formal statistical test is used in testing the ARCH effect. There are two available statistical tests for this; the Ljung-Box Q statistics and Langrange Multiplier (LM) test. If $a_t = b_t - \mu_t$ is the average corrected return, b_t is the return on the rate of exchange, μ_t is the average conditioned on b_t. The squared series b_t^2 detects the ARCH effects.

2.7.8 Ljung-Box Test

The test has the following hypothesis

H_o: data does not exhibit serial correlation or autocorrelation

H_1: data exhibits serial correlation

The test statistic is

$$Q = m \sum_{i=1}^{n} \hat{P}_i^2$$

$\hat{P}_i$ is the estimator of the autocorrelation component,

m is measurement of our sample under H_o,

the test statistic is assymptotically distributed as chi-square with n degrees of freedom.

For minute samples, the test statistic is

$$Q* = m(m+2)\sum_{i=1}^{n} \frac{\hat{P}_i^2}{T-i}$$

which also follows a chi-square distribution with m degrees of freedom under the H_o.

reject H_o if Q or Q* is larger than the corresponding critical value of the distribution with n degrees of freedom for a chosen significance level α.

2.7.9 Langrange Multiple (LM) Test

It is the same as the Fisher's F-statistic for experimenting $\beta_i = 0, (i = 1, ..., m)$ in the linear regression

$$x_t^2 = \beta_0 + \beta_1 x_{t-1}^2 + ... + \beta_n x_{t-n}^2 + \epsilon_t, t = n+1, ...M$$

ϵ_t is the nuisance term,

n is not a negative integer and

m is the measurement of the sample

$$H_o : \beta_1 = ... = \beta_n = 0$$

$$H_1 : \beta_1 \neq ... \neq \beta_n \neq 0$$

with test statistic

$$F = \frac{(SSR_o - SSR_1)/n}{SSR_1/(T-2n-1)}$$

$F \sim \chi_n^2$ under H_o where,

$$SSR_o = \sum_{t=m+1}^{T}(x_t^2 - \bar{M})^2,$$

$$\bar{M} = \sum_{t=1}^{T}\frac{x_i^2}{T}$$

is the sample mean of x_t^2 and

$$SSR_1 = \sum_{t=n+1}^{N}\hat{\epsilon}_t^2$$

$\hat{\epsilon}_t$ is the error component of the prior linear regression, reject the H_o if the test statistic of the chi-square distribution with n degrees of freedom is above the critical value for a defined significance level α or if the p-value of the distribution is smaller than the defined significance level (α).

2.7.10 Determination of the order of ARCH(p) model

When ARCH effect is confirmed in the data, the model is good and can be used to fit a series. Before that, the order (p) must be known. The partial autocorrelation function of the x_t^2 identifies the order p, of the ARCH (p) model.

Given that

$$x_t = \sigma_t\epsilon_t$$

$$\sigma_t^2 = \beta_0 + \beta_1 x_{t-1}^2 + ... + \beta_n x_{t-n}^2$$

with available sample, x_t^2 is an unbiased estimate of σ_t^2 and hence x_t^2 is projected to be linearly related to $x_{t-1}^2, ..., x_{t-n}^2$, which is same as an autoregressive model of order (p). A single x_t^2 is not a good estimate of σ_t^2, but rather serves as an approximate value that can give some idea about the order p.

2.7.11 Estimation of the ARCH(p) model

To estimate the ARCH (p) model, the error term is assumed to either follow normal distribution, standardized student t-distribution or generalized error distribution by using the likelihood. We shall assume the errors are normally

distributed.

Taking reference from the normality assumption associated with the error term ϵ_t, the likelihood function of an ARCH (p) model is given as

$$h(x_1, ..., x_t|\theta) = h(x_t|x_{t-1})h(x_{t-1}|x_{t-}$$

$$= \prod_{t=n+1}^{N} \frac{1}{\sqrt{(2\pi\sigma_t^2)}} exp(\frac{-x_t^2}{2\sigma_t^2}) f(x_1, ..., x_n|\theta) \quad 2.1$$

where $\theta = (\beta_o, \beta_1, ..., \beta_n)'$ and $h(x_1, ..., x_t|\theta)$ is the condensed form of probability density function of $x_1, ..., x_n$. Since the original expression of $h(x_1, ..., x_n|\theta)$ is complex and difficult to have, it is easily taken out from the prior likelihood function, particularly when the quantity of sample is large enough. Instead, it is simpler to condition on the first $x_1, ..., x_n$ because they are mostly identified and same as its observed values. This gives the conditional likelihood function:

$$h(x_1, ..., x_t|\theta; x_1, ..., x_n) = \prod_{t=n+1}^{N} \frac{1}{\sqrt{2\pi\sigma_t^2} exp(\frac{-x_t^2}{2\sigma_t^2})} \quad 2.2$$

where the calculation of σ_t^2 can be replicated. With reference from the normal distribution, the values of $\hat{\beta}_o, \hat{\beta}_1, ..., \hat{\beta}_n$ are calculated by maximising the prior likelihood function called the conditional maximum likelihood estimates(MLE) by Tsay (2006) as in (Sparks & Yurova, 2006).

Maximising the conditional likelihood function can be hard in handling. Another method which is much simpler is to maximise the logarithm of the conditional likeklihood function. Accordingly, the conditional log-likelihood function is:

$$l(x_{n+1}, ..., x_t|\theta; x_1, ...; x_m) = \sum_{t=n+1}^{N} (-\frac{1}{2} ln 2\pi - \frac{1}{2} ln\sigma_t^2 - \frac{x_t^2}{2\sigma_t^2})$$

$$= -\sum_{t=n+1}^{N}(2\sigma_t^2 + \frac{x_t^2}{2\sigma_t^2}) + K \tag{2.3}$$

where

$$K = \frac{-(T-n)}{2} ln(2\pi)$$

since the first term $\frac{1}{2} ln 2\pi$ has no parameter so taking it out does not affect the estimation. Again the calculation of $\sigma_t^2 = \beta_o + \beta_1 x_{t-1}^2 + ... + \beta_n x_{t-n}^2$ is replicated.

2.7.12 GARCH (p,q) Models

This type of model is a general form of the ARCH (p) model originally proposed by (Bollerslev, 1986). The ARCH (p) model violates the principle of parsimony in that it is over-parameterized with lots of (p) lags. The GARCH model has less parameter compared to the ARCH model hence making it a better choice. A special property of the GARCH is its usage of the lagged conditional variance terms as AR terms. It is given as:

$$x_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^{m} \beta_i x_{t-i}^2 + \sum_{j=1}^{n} \gamma_j \sigma_{t-j}^2$$

such that $x_t \sim N(0,1)$, and coefficients of the model are $\beta_i, i = 0, ..., k$ and $\gamma_j, j = 1, ...k$ such that $\beta_i \geq 0$ and $\gamma_j \geq 0$; $\sum_{i=1}^{max(m,n)}(\beta_i + \gamma_j) < 1$ and $\beta_i = 0$ for $i > m$ and $\gamma_j = 0$ for $j > n$. $\sum_{i=1}^{max(m,n)}(\beta_i + \gamma_j) < 1$ which ensures stationarity. $(\beta_i + \gamma_j)$ means the unconditional variance is limited in number but the conditional variance σ_t^2 relies on (m,n)

2.7.13 Estimating GARCH (p,q) model

After determining the orders of the GARCH model,we can proceed to find the coefficients of the model.In the process,we make use of the first values

of the squared returns and conditional variances of the past. For the initial values, the variances of the past associated with the returns can be used.Say $x_1, x_2, ..., x_n; \sigma_1^2, \sigma_2^2, ..., \sigma_s^2$ are identified, the conditional log-likelihood is given by

$$= l(x_{m+1}, ..., x_t; \sigma_{s+1}^2, ..., \sigma_t^2 | \theta; x_1, x_2, ..., x_m; \sigma_1^2, \sigma_2^2, ..., \sigma_s^2)$$

$$= \sum_{t=v+1}^{T} (-\frac{1}{2} ln(2\pi) - \frac{1}{2} In(\sigma_t^2) - \frac{x_t^2}{2\sigma_t^2}) \quad 2.4$$

where $\theta = (\alpha_o, \alpha_1, ..., \alpha_m; \beta_1, \beta_2, ..., \beta_s)$ and $v = max(p, q)$. Consequently, the conditional maximum likelihood estimates are achieved by maximizing the conditional log-likelihood function given the equation above.

2.7.14 Overview of Vector Autoregression Model

Vector Autoregressive (VARs) exceptionally concentrate on the effects of shocks. At the initial stage, it sorts out the essential shocks, then deductions from the calculation of impulse responses are portrayed. The present and past values of endogenous variables are used as a mechanism of forecasting in the VAR model.

Alternatively, individual endogeneous variables are described by their historical values and the historical values of any other endogenous variables in the VAR model. Moreover, it is much simpler to measure how relevant variables are when determining exchange rates using the variance decomposition analysis.

A key merit of the VAR is its ability to perform multiple time series analysis, since it provides a variety of requirements to set the efficient lengths of the variables. Significantly, VAR gives dual advantages. Firstly, it openly allows variables explain themselves hence considering mutual reliance between the variables involved. Secondly, compared to complex well-defined structural models, VAR analysis centers on brief relationship and just needs a simple model

with small number of variables to reach best efficiency.

A VAR model for a vector of macroeconomic variables is determined as follows:

$$\beta_0 y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + ... + \beta_p y_{t-p} + \mu_t$$

where

$y_t = (y_{1t}, y_{2t}, ..., y_{tk})'$ is a (k x 1) vector of macroeconomic variables,

$\alpha = (\alpha_1, ..., \alpha_k)'$ is a (k x 1) vector of coefficients matrix ; i = 1,2,..., k

p is the lag length

$\mu_t = (\mu_{1t}, \mu_{2t}, ..., \mu_{kt})'$ is a (kx1) vector of structured shocks

$$E(\mu_t) = 0$$
$$E(\mu_t \mu_t') = \sum_\mu = I_k$$
$$E(\mu_t \mu_s') = 0 \text{ for s}\neq\text{t}$$

$\sum_\mu$ is the covariance matrix and it is assumed to be non-singular. By definition, μ_t (structural shocks) are mutually uncorrelated, this means $\sum_\mu$ is a diagonal matrix.

To calculate the structural model, we first have to obtain the reduced form of equation above. This is achieved by pre-multiplying both sides of the equation by β_0^{-1}(the inverse of β_0).

$$\beta_0^{-1}\beta_0 Y_t = \beta_0^{-1}\alpha + \beta_0^{-1}\beta_1 Y_{t-1} + \beta_0^{-1}\beta_2 Y_{t-2} + ... + \beta_0^{-1}\beta_p Y_{t-p} + \beta_0^{-1}\mu_t$$

$$Y_t = \gamma + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + ... + \theta_p Y_{t-p} + \epsilon_t$$

where

$$\gamma = \beta_0^{-1}\alpha \text{ , } \theta_i = \beta_0^{-1}\beta_i \text{ , i=1,2,...,p and } \epsilon_t = \beta_0^{-1}\mu_t$$

is a white noise process with non-singular covariance matrix $\sum_{\epsilon}$

To transform from the reduced form to the structural model , we have to impose some identifying restriction. The coefficients $(\theta_{i's}, \gamma)$ of the reduced form in equation to the structural model in equation need to be restricted. From explanation above, $\sum_{\mu}$ (covariance matrix for μ_t) is taken to be diagonal, whereas β_0 has ones on its principal diagonal which is not restricted elsewhere. Meaning each element of Y_t has its own unique structural equation. The coefficients in the reduced form (γ and $\theta_{i's}$) in can be obtained using OLS. β_0 can be obtained by solving

$$\sum_{\epsilon} = cov(\epsilon_t) = cov(\beta_0^{-1}\mu_t) = \beta_0^{-1}\sum_{\mu}(\beta_0^{-1})'$$

$\sum_{\epsilon}$ contains K(K-1)/2 unique covariances (as a result of symmetry). Given that $\sum_{\epsilon}$ is diagonal and has K^2 elements, it means that K(K-1)/2 additional restrictions are needed to identify the system. When it comes to identification procedures in VAR models, we convert the reduced form in equation into a moving average format as shown below;

$$\theta(L)Y_t = \gamma + \epsilon_t$$

where

$\theta(L) = I - \theta_1 L - \theta_2 L^2 - ... - \theta_p L^p$ is the autoregressive lag order polynomial

If we assume that Y_t is a covariance stationary vector,the above equation can be represented in Wold moving average format as follows:

$$Y_t = \sum_{i=0}^{\infty} \phi_i \epsilon_{t-i} = \phi(L)\epsilon_t$$

where $\phi(L) = \theta(L)^{-1}$ and $\phi_o = I$

2.7.15 Model Specification

In order to examine the level of impact macroeconomic indicators have on exchange rate systems, the study used the Vector Autoregressive model by (Sims, 1980). Single variable Autoregression is a simple equation, variables in this singular model are described by past values of itself.

VAR on the other hand is a model that possess n-variable equation, where each variable taken is described by past values of itself together with past values of the remaining variables in the model. It gives a step-by-step approach in observing valuable transformations and its statistical analysis are simpler to adopt and explain.

Likewise Sims (1980) and others in lots of impactful research papers concluded that VARs offer reliable steps in portraying data as well as better policy analysis.

$$\text{EXRATE} = \text{f(GDP,INF)}$$

$$\text{GDP} = \text{f(EXRATE,INF)}$$

$$\text{INF=f(EXRATE,GDP)}$$

Transforming into statistical model as

$$\text{EXRATE} = \beta_o + \beta_1 GDP + \beta_2 INF + \epsilon_t$$

$$\text{INF=}\beta_o + \beta_1 EXRATE + \beta_2 GDP + \epsilon_t$$

$$\text{GDP =}\beta_o + \beta_1 EXRATE + \beta_2 INF + \epsilon_t$$

Inflation rate (INF), GDP and Exchange rate (EXRATE)

2.7.15.1 Explanation of Variables

2.7.15.2 GDP

The constant price is calculated as an output of the economy. It is not affected by adjustments in price level. Nominal GDP on the other hand are mostly disturbed

by volatilities leading to poorer representation (Mankiw & Reis, 2006). Real GDP gives a better representation of output. Whiles GDP increases, exchange rate decreases. Hence an inverse relationship between the two (Khordehfrosh Dilmaghani & Tehranchian, 2015; Muchiri, 2017). β_1 is expected to be negative. $\beta_1 < 0$

2.7.16 Inflation Rate

There are lots of ways to measure inflation, some are the GDP deflator or CPI. The CPI is mostly used in Ghana.CPI was used in (Mankiw & Reis, 2006). It was proven that there is a direct correlation between inflation and exchange rate (Khordehfrosh Dilmaghani & Tehranchian, 2015; Okoth, 2013).

2.7.16.1 Exchange Rate

The rate at which one currency is substituted for the other is termed the nominal exchange rate. The nominal exchange rate does not show if there is a change in the value of a local currency alongside its trading partners without the US dollar. It is silent on the impact of accessing foreign goods and services on the rate itself (Appleyard & Suresh, 2016).

2.7.17 Granger-Causality Test

A component under study is said to Granger-cause another variable if the former helps in determining the later or if the numeric term of the lagged values of its former are statistically significant. The Granger causality test by Granger (1969) as cited in Emmanuel (2015) is determined by how far to which the variable under investigation for instance Y can be described by a past value for instance W and also to examine whether including past values can enhance the description.

The Granger causality test must be used on just stationary variables (I(0)) series and to confirm whether δW helps in determining δY, F-test is employed. The

test is depicted as ;

$$\delta W_t = \prod_i^n = 1\alpha_1 i \delta W_{t-i} + \prod_i^n = 1\Psi_1 i \delta Y_{t-i} + \epsilon_1 t$$

$$\delta Y_t = \prod_i^n = 1\alpha_2 i \delta W_{t-i} + \prod_i^n = 1\Psi_2 i \delta Y_{t-i} + \epsilon_2 t$$

The null hypothesis is W does not Granger-cause Y, and Y does not Granger-cause W. Thus

$$H_o : \Psi_1 i = 0 \text{ and } H_o : \alpha_2 i = 0 \ , \quad i = 1, 2, ..., \ m.$$

The two variables are distinct if H_o is not rejected meaning there no causal relationship between them. When both H_o are rejected it means that there is causality in both direction and if just one is rejected, it means a unidirectional causality among the variables. The Granger causality test is adopted since the study wants to figure out causality between exchange rate and the other variables.

2.7.17.1 Priori Expectation

When it comes to the anticipated outcomes of the variable or component added to our model after a positive exchange rate shock or turbulence as well as turbulences to other specified components, the variables are projected to react as shown in Table (2.1) below. Prices represented by inflation is projected to decrease after the shock.

For GDP, the direction of the outcome hugely depends on the degree of shortrun non-neutrality of money. However, GDP is expected to decrease after the shock. Since Ghana is an inflation targeting country, we project a high impact on the outcome of inflation than on GDP since the main goal of inflation targeting is price stability.

For the exchange rate, we project that a rigid policy would lead to an increase in interest rate, that would make the cedi gain value (appreciate). If however,

a rigid policy results in a decrease in the exchange rate, then the cedi will lose value (depreciate). Thus the outcome of exchange rate relies on the Fisher effect.

Table 2.1: A priori expectation of the variables

	Response	
Shock	GDP	Inflation
Exchange rate	Rise	Rise

2.7.18 Background on Bayesian Statistics

Mathematical statistics is centered on two main concepts: frequentist and Bayesian statistics.

Bayesian statistics has a long history with much controvesy dating back to the 1700's. A history on Bayes rule and how it was used in locating a lost submarine can be found in (McGrayne, 2011).

To start with the main distinction among the two concepts, a clear difference between the two is that Bayesian statistics involves prior knowlegde and observed data where frequentist statistics only involves observed data. Using prior observation can be subjective, making it a field of criticism in statistics.

Next, there are distinction in the nature of probability, parameters and inferences. O'Hagan and Luce (2003) as cited in Spiegelhalter (2004) highlighted some brief differences between the two methods.

Also, how inferences are done are different. Frequentist statistics uses a confidence interval where Bayesian statistics uses a credible interval. Confidence intervals are like frequency statements. It takes random samples from a

population and concludes that some percent of the samples contain the true parameter. A credible interval is a probability statement. It conlcudes that there is some percent chance that the interval contains the true parameter.

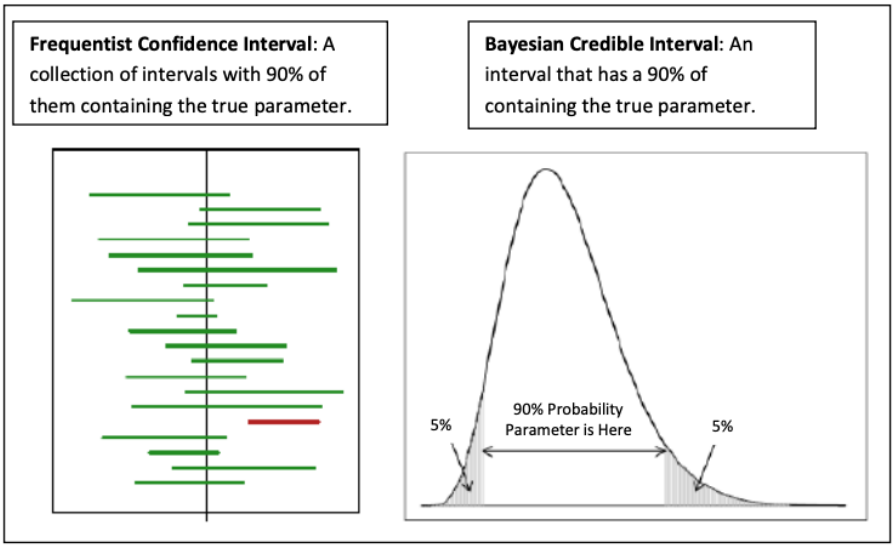
Summary of Key Differences between Frequentist and Bayesian Statistics (O'Hagan and Luce, 2003)

FREQUENTIST	BAYESIAN
Nature of probability	
Probability is a limiting, long-run frequency.	Probability measures a personal degree of belief.
It only applies to events that are (at least in principle) repeatable.	It applies to any event or proposition about which we are uncertain.
Nature of parameters	
Parameters are not repeatable or random.	Parameters are unknown.
They are therefore not random variables, but fixed (unknown) quantities.	They are therefore random variables.
Nature of inference	
Does not (although it appears to) make statements about parameters.	Makes direct probability statements about parameters.
Interpreted in terms of long-run repetition.	Interpreted in terms of evidence from the observed data.
Example	
"We reject this hypothesis at the 5% level of significance."	"The probability that this hypothesis is true is 0.05."
In 5% of samples where the hypothesis is true it will be rejected (but nothing is stated about this sample).	The statement applies on the basis of *this* sample (as a degree of belief).

From the above illustration, it is obvious that there are differences between the Bayesian and frequentist methods of making inferences in statistics.

Nevertheless, eventhough some controversies existed behind the above mathematical inferences, the aim of this study is to develop a method useful in economic application in Ghana. Bayesian statistics has been applied in many fields and can also be useful in forecasting macroeconomic variables.

The Bayes concept of making inferences is discussed as follows; The goal of the Bayesian model is to evaluate posteriors that helps in calculating an unknown



Confidence interval versus Credible interval (Recchia, 2012)

statistic based on a likelihood function and a determined prior distribution.

Suppose the statistic of interest is exchange rate (θ) and its unknown and random as well. The prior distribution $P(\theta)$ stands for any prior knowledge we know about exchange rate before observing it. This distribution can be "non-informative" and objective, where the prior does not favour any value of θ, or it can be "informative" and subjective where the prior is formed based on expert opinion. The likelihood function $P(y|\theta)$ is the term given to the model being applied to the observed data.

The posterior $P(\theta|y)$ is the end result that enables us to calculate what we believe is the true exchange rate based on prior knowledge (prior distribution) and observed data (likelihood function). The prior likelihood and posterior are all related in Bayes' rule;

$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{P(y)} = \frac{P(y|\theta)P(\theta)}{\int P(y|\theta')P(\theta')}$$

The next step changes $P(y)$ using the law of total probability. The challenge is that the integral in the denominator is difficult to manage making Bayesian analysis very complex .

However, there are procedures to avoid this integral. The procedures are either;

- The use of a conjugate prior
- The use of a hierarchical prior
- The use of a Monte Carlo simulation

In brief, the Bayesian statistics differs from frequentist statistics which has been the subject of controversy when using Bayes' rule (Bergman et al., 2017). However, it has been used successfully in practice especially for problems with limited data.

Despite the fact that the VAR model is suitable for macroeconomic analysis and adopted by many researchers, the over-parameterizaton problem of the model violates the parsimonious rule of modelling. For a model to be adjudged as best, it has to be simple with less forecasting errors. This study makes use of the Bayesian VAR model which is also adopted in macroeconomic analysis, but has not gained prominence in the analysis of macroeconomic variables in Ghana.

2.8 Summary

It is obvious from the literature that the use of the Bayesian VAR in forecasting exchange rates in Ghana is absent. Meanwhile, it has been used in other parts of the world in analysing different data sets. Most of the literature used the prior by (Litterman, 1986).

Hence, this study will use the Bayesian VAR in projecting the exchange rate in Ghana using the Multivariate Normal Distribution as the prior. This is further discussed in Chapter 3.

Chapter 3

METHODOLOGY

3.1 Introduction

This chapter touched on the methodology for the study and it has been classified into sections. The sections covered topics such as the research design, the data and data source, stationarity tests, the ARIMA model, the EGARCH model and the Bayesian VAR model.

3.2 Research Design

The study employed a time series approach which was mainly quantitative and the development of mathematical models to examine the connection among quantitative measurements.

3.3 Source of Data

The study used monthly data on exchange rate only for the analysis of the volatility since the ARCH-family models were univariate time series models. Inflation, exchange rate and GDP were used for the Bayesian Analysis. The data spanned from January, 2008 to March, 2022. The data was obtained from the Bank of Ghana's official website.

The data was segmented into training and testing sets. The training set was used in the estimation of the model's parameters while the testing set was used to check the accuracy of the model's forecast values.

3.4 The Exchange Data

Monthly data of exchange rate between the cedi and the dollar from the Central Bank was used for the study. It contained monthly data on exchange rate from January, 2008 to March, 2022. 171 observations of the data were realized. The Bayesian aspect of the analysis considered data on other factors (Inflation and GDP) that impacted the uncertain nature of the exchange rate.

In this work, returns (b_t) were observed as continuously accumulated returns which were the first difference in logarithm of the Interbank exchange rate.

$$b_t = ln(\frac{a_t}{a_{t-1}}) \tag{3.1}$$

Where a_t was the cedi to dollar exchange rate at time t and a_{t-1} was the exchange rate at time (t-1). The b_t in equation (3.1) was used in examining the fluctuations of the Interbank exchange rate.

3.5 Stationarity Tests (Unit Root Test)

Time Series can only be performed when the data is stationary. Thus, the first obligation before starting a time series analysis is to check if the series exhibits stationarity. Mostly, stationarity has to be established.

A series is said to be stationary if the mean and variance do not exhibit any trend with respect to time, or do not vary with time. It is further classified into two, that is, strict stationarity and weak stationarity.

Strict stationarity has joint distribution function that is not influenced by

changes in time. Say a series y_t is strictly stationary if the joint distribution $(y_{t1}, y_{t2}, ..., y_{tk})$ is identical to that of $(y_{t1+t}, ..., y_{tk+t})$ for all t, where k is a positive integer and $(t_1, ..., t_k)$ is a collection of k positive integers. Shifting or changes in time (t) does not affect the joint distribution which only relies on the intervals given by t which is called a lag. This type of stationarity is not common in time series because it is difficult to apply hence weak stationarity is mostly considered (Akumbobe, 2015).

A series y_t is weakly stationary if its first and second moments are independent of time.

- $E(y_t) = \mu$ (a constant)
- $Cov(y_t, y_{t-k}) = \gamma_k$ (depends only on its lag(difference))

This study tried to find out if the variables of interest possess some properties of time series to enable modelling to be done on the data. Procedures need to be followed to verify if the variables are stationary or not. Most at times, stochastic trends are observed in a lot of time series data (Naceur et al., 2007). The availability of a stochastic trend is known by testing the availability of unit roots (Dickey & Fuller, 1979; Dickey & Fuller, 1981; Phillips & Perron, 1988; Kwiatkowski et al., 1992).

There are lots of methods for testing if a time series is stationary or not. Examples are the Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) test, Phillip-Perron test, Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test and NP test.

This study used the ADF and KPSS test to verify if the series is stationary and to confirm the existence of unit roots in the study.

The Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and the Hannan-Quinn (H-Q) Information Criterion are used to choose

the lag length for both the ADF and the KPSS test. The ADF is specified as follows:

$$\Delta X_t = \alpha + \delta_t + \rho X_{t-1} + \sum_{i-1}^{\rho} \beta_i \Delta X_{i-1} + V_i \tag{3.2}$$

where,

X_t denotes the time series at time t

Δ is the difference operator

α,δ,β are parameters that shall be estimated and

V is the stochastic random distribution term.

The KPSS test is more powerful compared to the ADF test because of the following reasons. Firstly, the ADF test does not account for heteroscedasticity and non-normality which are mostly available in raw econometric time series data.

Secondly, ADF cannot decide between stationary and non-stationary series that can cause autocorrelation. Again, in cases where the variables under investigation exhibit serial correlation and structural breaks.

The hypothesis tested in KPSS tests is as follows :

H_o : Series is level stationary against

H_1: Series is not level stationary

while that of ADF test is :

H_o: Series contain unit root against

H_1:Series does not contain unit root

From the ADF test, the null hypothesis means non-stationarity against the alternative hypothesis, that is it has no unit root meaning stationary.

We reject the null hypotheses of the ADF test and conclude that there is no unit root meaning stationarity, **when the p-value is smaller than the critical value** and we fail to reject the null hypotheses of the KPSS test and conclude that the series is level stationary when **when the p-value is greater than the critical value**. The reverse is also true.

3.6 ARIMA Models

ARIMA models joins together the differenced series of past values, thus the "I" part being explained by the Autoregressive (AR) part, the residual terms of the past values explained by the Moving Average (MA) part.

That is an ARMA transforms into ARIMA if the series is stationary after being differenced. ARIMA (p,d,q) means that there are 'p' number of lagged or historical values of the series, "d" number of times series has been differenced for it to be stationary, and "q" number of lagged or historical values of the residual parts.

Meaning that, if ARMA (p,q) transforms into ARIMA(p,d,q) that would be interpreted as

$$\Delta^d Y_t = (1 - B)^d Y_t \tag{3.3}$$

"d" standing for the counts of the differencing on the series before stationarity. Y_t refers to the series itself and B is the lag order of the series. Usually, differencing can be done at most two times for normal data. Excessive differencing brings problems such as missing critical observations or dynamics in the data. The ARIMA model can thus be generally represented as;

$$\phi(B)(1 - B)^d Y_t = \phi(B)\epsilon_t \tag{3.4}$$

And if $E(\delta^d)Y_t = \mu_t$, then the model is represented as;

$$\phi(B)(1-B)^d Y_t = \alpha + \theta(B)\epsilon_t \quad 3.5$$

Describing α as $\alpha = \mu(1 - \phi_1 - ... - \phi_p)$

3.6.1 Assessment of Univariate Time series Models

Estimating and forecasting of univariate time series models is executed with the Box-Jenkins methodology. It refers to the set of procedures with the following steps;

- Identification
- Estimation
- Diagnostic checking

The Box-Jenkins methodology is applicable only to stationary variables and so the first thing to observe before modelling or carrying out the steps as listed above is to check for stationarity.

3.6.1.1 Model Identification

A preliminary Box-Jenkins analysis with a plot of the initial data should be run as the starting point in determining an appropriate model. The identification approach is to first determine the order of differencing, d , neccessary to induce stationarity and then the autoregressive order p, and the moving average order q.

At identification stage of a data, two tools are used to measure the correlation between the observations within a single series of data. These tools are the autocorrelation function (ACF) and the partial autocorrelation function (PACF). They are diagrams that identify the possible orders of both AR and MA to be

selected.

AR (p) means the ACF tails of as exponential decay or damped siine wave whiles the PACF cuts off after lag p, MA (q) means the ACF cuts of at lag q and the PACF tails of as exponential decay or damped sine wave and ARMA (p,d) means the ACF tails off after lag (q-p) and the PACF tails off after lag (q-p).

A summary of the characteristics of the theoretical ACF and PACF is demonstrated below.

Table 3.1: The ACF and PACF theoretical behaviour for model identification

Model	ACF	PACF
AR (p)	spikes decays to zero	cutt off after lag q
MA (q)	Cut off after lag p	spikes decays to zero
ARIMA (p,d)	spikes decays to zero	spikes decays to zero

3.6.1.2 Maximum Likelihood Estimation of ARMA models

For independent and identically distributed (iid) data with marginal pdf $f(y_t;\theta)$, the joint density function for a sample $y = (y_1, ..., y_T)$ is simply the product of the marginal densities for each observation.

$$f(y;\theta) = f(y_1, ..., y_T;\theta) = \prod_{t=1}^{T} f(y_t;\theta) \tag{3.6}$$

The likelihood function is this joint density treated as a function of the parameter θ given the data y:

$$L(\theta|y) = L(\theta|y_1, ..., y_T) = \prod_{t=1}^{T} f(y_t;\theta) \tag{3.7}$$

The log-likelihood then has the simple form

$$\ln L(\theta|y) = \Sigma_{t=1}^{T} \ln f(y_t;\theta) \tag{3.8}$$

For a sample from a covariance stationary time series y_t, the construction of the log-likelihood given above does not work becuase the random variables in the sample $(y_1, ..., y_T)$ are not independent and identically distributed.

One way is to rely on factorization of the joint density into a series of conditional densities and the density of two adjacent observations $f(y_2, y_1;\theta)$ from a covariance stationary time series. The joint density can always be factored as the product of the conditional density of y_2 given y_1 and the marginal density of y_1;

$$f(y_2, y_1;\theta) = f(y_2|y_1;\theta)f(y_1;\theta) \tag{3.9}$$

For three observations, the factorization

$$f(y_3, y_2, y_1;\theta) = f(y_3|y_2, y_1;\theta)f(y_2|y_1;\theta)f(y_1;\theta) \tag{3.10}$$

In general, the conditional marginal factorization has the form

$$f(y_T, ..., y_1;\theta) = (\prod_{t=p+1}^{T} f(y_t|I_{t-1},\theta)).f(y_p, ..., y_1;\theta) \tag{3.11}$$

where $I_t = y_t, ..., y_1$ denotes the information available at time t, and $y_p, ..., y_1$ denotes the initial values. The log-likelihood function may then be expressed as

$$\ln L(\theta|y) = \sum_{t=p+1}^{T} \ln f(y_t|I_{t-1},\theta) + \ln f(y_p, ..., y_1;\theta) \tag{3.12}$$

The full log-likelihood function is called the exact log-likelihood. The first term is called the conditional log-likelihood, and the second term is called the marginal log-likelihood for the initial values.

In the MLE of time series models, two types of MLE's may be computed. The first type is based on maximizing the conditional log-likelihood function. These estimates are called conditional MLE's and are defined by

$$\hat{\theta}_{cmle} = argmax_{\theta} \sum_{t=p+1}^{T} \ln f(y_t|I_{t-1}, \theta) \tag{3.13}$$

The second type is based on maximizing the exact log-likelihood function.These estimates are called exact MLE's and are defined by

$$\hat{\theta}_{mle} = argmax_{\theta} \sum_{t=p+1}^{T} \ln f(y_t|I_{t-1}, \theta) + \ln f(y_p, ..., y_1; \theta) \tag{3.14}$$

For stationary models, $\hat{\theta}_{cmle}$ and $\hat{\theta}_{mle}$ are consistent and have the same limiting normal distribution. In finite samples, however, $\hat{\theta}_{cmle}$ and $\hat{\theta}_{mle}$ are generally not equal and may differ by a substantial amount if the data are close to being non-stationary or non-invertible.

For example,for an MLE of an AR(1) model

Consider the stationary AR(1) model

$$y_t = c + \psi y_{t-1} + \epsilon_t, \quad \epsilon \sim N(0, \sigma^2), \quad t = 1, ..., T \tag{3.15}$$

$$\theta = (c, \psi, \sigma^2)', \quad |\psi| < 1 \quad \text{conditional on} \quad I_{t-1}$$

$$y_t|I_{t-1} \sim N(c + \psi y_{t-1}, \sigma^2), \quad t = 2, ..., T$$

which only depends on y_{t-1}. The conditional density $f(y_t|I_{t-1}, \theta)$ is then

$$f(y_t|y_{t-1}, \theta) = (2\pi\sigma^2)^{-\frac{1}{2}} exp(-\tfrac{1}{2\sigma^2}(y_t - c - \psi y_{t-1})^2), \quad t = 2, ..., T$$

To determine the marginal density for the initial values y_1, recall that for a statioanry AR(1) process

$$E[y_1] = \mu = \frac{c}{1-\psi}$$

$$var(y_1) = \frac{\sigma^2}{1-\psi^2}$$

It follows that

$$y_1 \sim N(\frac{c}{1-\psi}, \frac{\sigma^2}{1-\psi^2})$$

$$f(y_1;\theta) = (2\pi\frac{\sigma^2}{1-\psi^2})^{-\frac{1}{2}} exp(\frac{-1-\psi^2}{2\sigma^2}(y_1 - \frac{c}{1-\psi})^2)$$

The conditional log-likelihood function is

$$\sum_{t=2}^{T} \ln f(y_t|y_{t-1},\theta) = \frac{-(T-1)}{2}\ln 2\pi - \frac{T-1}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=2}^{T}(y_t - c - \psi y_{t-1})$$

We notice that the conditional log-likelihood function has the form of the log-likelihood function. It follows that the conditional MLEs for c and ψ are identical to the least squares estimate from

$$y_t = c + \psi y_{t-1} + \epsilon_t, t = 2, ..., T$$

and the conditional MLE for σ^2 is

$$\hat{\sigma^2}_{cmle} = (T-1)^{-1}\sum_{t=2}^{T}(y_t - \hat{c}_{cmle} - \hat{\psi}_{cmle}y_{t-1})^2$$

The marginal log-likelihood for the initial value y_1 is

$$\ln f(y_1;\theta) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln(\frac{\sigma^2}{1-\psi^2}) - \frac{1-\sigma^2}{2\sigma^2}(y_1 - \frac{c}{1-\psi})^2$$

The exact log-likelihood function is then

$$\ln L(\theta|y = -\frac{T}{2}\ln 2\pi - \frac{1}{2}\ln\frac{\sigma^2}{1-\psi^2} - \frac{1-\psi^2}{2\sigma^2}(y_1 - \frac{c}{1-\psi})^2 \ -$$

$$\frac{T-1}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=2}^{T}(y_t - c - \psi y_{t-1})^2 \qquad 3.16$$

The exact log-likelihood function is a non-linear function of the parameter θ, and so there is no closed form solution for the exact MLEs. The exact MLE must be determined by numerically maximizing the exact log-likelihood function usually a Newton-Raphson method is used for the maximization which leads to the iteractive scheme

$$\hat{\theta}_{mle,n} = \hat{\theta}_{mle,n-1} - \hat{H}(\hat{\theta}_{mle,n-1})^{-1}\hat{S}(\hat{\theta}_{mle,n-1})$$

where $\hat{H}(\hat{\theta})$ is an estimate of the Hessian matrix (2nd derivative of the log-likelihood function), $\hat{S}(\hat{\theta})$ is an estimate of the score vector (1st derivative of the log-likelihood function). The estimates of the Hessian and score may be computed numerically (using numerical derivative routines) or they may be computed analytically (if analytic derivatives are known).

3.6.1.3 Maximun Likelihood Estimation of MA(q)

If $\epsilon = (\epsilon_1, ..., \epsilon_T)'$

Then $\epsilon = NX + Z\epsilon^*$

$$= \begin{pmatrix} Z & N \end{pmatrix} \begin{pmatrix} \epsilon^* \\ X \end{pmatrix}$$

with

$$X = \begin{pmatrix} x_1, ..., x_T \end{pmatrix}', \quad \epsilon^* = \begin{pmatrix} \epsilon_{1-q}, \epsilon_{2-q}, ..., \epsilon_{-1}, \epsilon_o \end{pmatrix}'$$

and

$$\text{N}=\begin{pmatrix} O_{qxT} \\ \\ A_1(\theta) \end{pmatrix}, \quad \text{Z} = \begin{pmatrix} I_q \\ \\ A_2(\theta) \end{pmatrix}$$

where $A_1(\theta)$ is a lower triangular matrix (T x T), $A_2(\theta)$ is a T x q matrix and I_q is the identity matrix of order q.

The joint pdf of ϵ is defined to be

$$(2\pi\sigma_\epsilon^2)^{\frac{T+q}{2}} exp[-\tfrac{1}{2\sigma_\epsilon^2}(NX+Z\epsilon^*)'(NX+Z\epsilon^*)]$$

The marginal density of X is defined to be

$$(2\pi\sigma_\epsilon^2)^{-\frac{T}{2}}(detX'X)^{-\frac{1}{2}} exp[-\tfrac{1}{2\sigma_\epsilon^2}(NX+Z\hat{\epsilon}_o)'(NX+Z\hat{\epsilon^*})]$$

with

$$\hat{\epsilon^*} = (Z'Z)^{-1}Z'NX$$

The log-likelihood function is defined to be

$$l(x;\sigma) = -\tfrac{T}{2}log(2\pi) - \tfrac{T}{2}log(\sigma_\epsilon^2 - \tfrac{T}{2}log(detX'X) - \tfrac{S(\theta)}{2\sigma_\epsilon^2}$$

where

$$\sigma = (\theta', \sigma_\epsilon^2)' \text{ and } S(\theta) = (NX+Z\hat{\epsilon^*})'(NX+Z\hat{\epsilon^*})$$

Using the first-order condition with respect to σ_ϵ^2 one has

$$\hat{\sigma}^2_{\epsilon,ml} = \tfrac{S(\hat{\theta}_{ml})}{T}$$

The concentrated log-likelihood function is defined to be

$$l^*(x;\theta) = -Tlog[S(\theta)] - log[detX'X]$$

Therefore $\hat{\theta}_{ml}$ gives $min_\theta[Tlog[S(\theta)] + log[detX'X]]$

3.7 Volatility models

3.7.1 The Exponential GARCH (p,q) Model

This model was introduced by Nelson (1991) to correct the limitations of the GARCH model. The model accounts for the assymmetric effects, that is, positive and negative shocks. EGARCH is able to remove the limitations on the parameters. The non-negativity of parameters which is a problem of the GARCH is dealt with properly .

Negative shocks results in large fluctuations than positive shocks hence causes more damage (Zivot, 2009). In this type of models, upward and downward trends of shocks are called good and bad news. If a fall of a return is followed by a huge fluctuation which is larger than the fluctuations caused by the rise then it is called a leverage effect.

The EGARCH can be specified as

$$x_t = \sigma_t \epsilon_t$$

$$\ln \sigma_t^2 = \alpha_0 + \frac{\sum \alpha_i |x_{t-i}| + \gamma_i x_{t-i}}{\sigma_{i-1}} + \sum \beta_j \ln(\sigma_{t-j}^2) \qquad 3.17$$

Positive shocks x_{t-i} results in an effect $\alpha_i(1+\gamma_i)|\epsilon_{t-i}|$ on the logarithm of the conditonal variance while negative shocks x_{t-i} results in an effect of $\alpha_i(1-\gamma_i)|\epsilon_{t-i}|$, where $\epsilon_{t-i} = \frac{x_{t-i}}{\sigma_{i-1}}$.γ represents the leverage effect and it has to be negative. Normally, $\alpha_i < 0$,$\gamma_i < 0$, also , $\beta_i + \gamma_i < 1$

Taking the logarithm of the rates makes sure that the model reacts to both negative and positive lagged values of x_t asymmetrically.

3.7.2 Estimation of EGARCH model parameters

Let

$$y_t = f(w_t; \rho) + \epsilon, t = 1, ..., T \tag{3.18}$$

where f is at least twice continuously differentiable function of ρ, with $w_t = (1, y_{t-1}, ..., y_{t-n}, x_{1t}, ..., x_{kt})'$. The error process is parameterized as

$$\epsilon_t = z_t h_t^{\frac{1}{2}}, t = 1, ..., T$$

By assuming normality, the log-likelihood function of the EGARCH (p,q) model is

$$L_t = c - \frac{1}{2}\sum_{t=1}^{T} \ln h_t - (\frac{1}{2})\sum_{t=1}^{T}(\epsilon_t^2 | h_t) \tag{3.19}$$

with

$$\ln h_t = \alpha_o + \sum_{j=1}^{q} \alpha_j Z_{t-j} + \psi(|Z_{t-j}| - E|z_t|) + \sum_{j=1}^{p} \beta_j \ln h_{t-j} \tag{3.20}$$

Let $\beta = (\alpha_o, \alpha_1, ..., \alpha_q, \psi_1, ..., \psi_q, \beta_1, ..., \beta_p)'$.

The first partial derivatives with respect to the EGARCH parameters are

$$\sum_{t=1}^{T} \frac{\partial l_t}{\partial \beta} = \frac{1}{2}\sum_{t=1}^{T}(\frac{\epsilon_t^2}{h_t} - 1)\frac{\partial \ln h_t}{\partial \beta} \tag{3.21}$$

where

$$\frac{\partial \ln h_t}{\partial \beta} = X_{\beta t} + (\tfrac{1}{2})\textstyle\sum_{j=1}^{q} \alpha_j Z_{t-j} + \psi_j |Z_{t-j}| \frac{\partial \ln h_{t-j}}{\partial \beta} + \sum_{j=1}^{p} \beta_j \frac{\partial \ln h_{t-j}}{\partial \beta}$$

and

$$X_{\beta t} = (1, Z_{t-1}, ..., Z_{t-q}, |Z_{t-1}| - E|Z_t|, ..., |Z_{t-q}| - E|Z_t|, \ln h_{t-1}, ..., \ln h_{t-p})'$$

The parameters of equation (3.18) with EGARCH errors of equation (3.28) may be estimated jointly by maximum likelihood. The normality assumption

guarantees block diagonality of the information matrix such that the off-diagonal blocks involving partial derivatives with respect to both mean and variance parameters are null matrices.

Thus the parameters of the conditional mean defined by equation (3.18) can be estimated without asymptotic loss of efficiency. This implies that maximum likelihood estimates for the parameters in equation (3.28) can be obtained numerically from the first-order conditions defined by equation (3.21) to zero.

3.7.3 Evaluation of EGARCH(1,1) model

Using the alternative form of the EGARCH (p,q) model, the EGARCH (1,1) can be written as

$$x_t = \sigma_t \epsilon_t$$

$$\ln \sigma_t^2 = \alpha_o + \frac{\alpha_1 |x_{t-1}| + \gamma_1 x_{t-1}}{\sigma_{t-1}} + \beta_1 \ln \sigma_{t-1}^2$$

3.7.4 Forecasting with EGARCH Model

The projection done with the EGARCH model is the same as that of the ARMA model because it uses the ARMA properties to elaborate the discovery of the conditional variance of x_t.

3.8 The Model Selection (Information) Criteria

For us to find a better model to fit our data set, there are lots of models available. We can then use different measurement methods to find the best model. The Autocorrelation and Partial Autocorrelation functions helps in determining the rank of our model, however it gives us a rough idea on the

construction of our model based on our proposed rank of the model (Aidoo, 2010).

One of the methods one can use is the information criterion test. The information criteria gives a metric for weighing information using the order of the model. It combines test on the goodness of fit and parsimony determination of our model.

One theory developed for this information criterion test is to discover the best model with least number of coefficients. We do so by using the MLE for each model.

We shall consider the Aikaike Information Criteria (AIC) by Akaike (1974), Bayesian Information Criterion (BIC) by Schwartz (1978) and Hannan-Quinn information criteria (H-Q) test by (Hannan & Quinn, 1979). The AIC is used as the measure of the goodness of fit test. It is the most commonly used and there has been lots of works done on it. The AICc is an extension of the AIC with a second order correction for small sample sizes (Akaike, 1974).

The best model is selected based on their ranking (order) by these information criteria. The model with the least AIC ,AICc , BIC, and H-Q is chosen as the best model among other models.

If any two or more of the models have equal AIC, BIC or H-Q, the parsimony rule is applied to choose the right model. Otherwise, we can use the loss functions of the models involved to select (Aidoo, 2010).

In the general case, the AIC, BIC , AICc and H-Q are calculated as;

$$\begin{aligned} AIC &= -2lnl + 2k \\ AICc &= AIC + \frac{2k(k+1)}{T-k-1} \\ HQ &= -l + 2kln(InT) \\ BIC &= -2lnl + kIn(T) \end{aligned} \tag{3.22}$$

where;

l= the maximized value of the log likelihood function

k= number of regressors/parameters used in the model

T= number of observations or the sample size

3.8.1 Diagnostic Checking

We employ model diagnostics to enable us check the performance of a model to detect its adequacy or goodness of fit. We conduct this on the residuals especially on the standardized residuals (Sammy, 2018). The residuals or error terms are independently and identically distributed normal distribution (Tsay, 2005). Histogram and quantile plot of the errors are some of the plots that can be applied.

If our model perfectly match the data, our histogram of the errors will depict a symmetric bell-shape. That is the quantile plot will have most of the observations from the data on the straight line, the quantile plot matches our sample observations to the best fitted probability distribution generally (Suhartono, 2005).

If our model meets all model requirements stated including relevant statistical tests, then it is accepted as the right mathematical representation of the data. With these established, we can proceed to use it in forecasting.

3.9 Model Evaluation

The data was put into two parts. The first part being the training set and the other part the testing set. The training set was used to estimate the model coefficients whilst the testing set was used to validate the model.

This process is very crucial to estimate the model to know how precise it is in forecasting. If our chosen model is able to give a good explanation of the testing set, then we call our model valid and adequate, which can be used in forecasting.

3.10 Forecasting Accuracy of the Model

As said earlier, forecast accuracy test is another method for choosing the ultimate model. There are lots of ways of assessing the forecast accuracy of models. Some are the: Mean Square Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and the Theil's U-Statistic. MSE is the mean of the squared difference between the actual variance and the volatility forecast (σ_t^2). When the true variance obtained is absent, the squared series observation x_t^2 is applied.The MSE is given by:

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

where $\hat{\sigma_t^2}, t = 1, ..., T$ is the approximated conditional variance achieved from fitting the ARCH-type models. The MSE is looked down upon since x_t^2 is noisy and unstable although the x_t is a stable estimator of σ_t^2 (Tsay, 2005). Alternatively, other methods have been suggested. Lopez (2001) proposed the MAE as

$$MAE = \frac{\sum_{t=1}^{T}|x_t^2 - \hat{\sigma_t^2}|^2}{T}$$

3.10.1 Root Mean Squared Error (RMSE)

It is an extension of the MSE and it is usually used by researchers for assessing models. It is able to improve upon the estimates of the MSE. The RMSE is given as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$$

3.10.2 Mean Absolute Percentage Error (MAPE)

It is also one of the common measures of accuracy in most research works. Even though, it is simple in usage, it has a hindrance of not being able to estimate if there were zero's available in the data. The MAPE is given as:

$$MAPE = \frac{100\%}{n} = \sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right|$$

3.11 Bayesian Statistics

Statistical inference is about making deductions based on scientific evidence from data. There are several ways of making deductions examples are Statistical modelling, data oriented strategies and uncensored use of designs and randomization in analysis as in (Tutuani, 2015).

There are lots of rules such as the Bayesian, for undertaking statistical deductions. The Frequentist deduction uses frequency translation of probability, and the Bayesian deduction is centered on the level of belief translation.

As a result, the Bayesian approach to making deductions is a way of making statistical deductions where Bayes' Theorem is applied to keep up-to-date the probability of a hypotheses as enough proof is provided.

This procedure which derived its name from Thomas Bayes (1701-1761), is a theory in statistics where proof about the nature of the universe is explained in relation to the levels of belief called Bayesian probabilities. This is a particular type from a bouquet of methods of probability and there are several statistical methods that do not center on levels of belief.

Edwards et al. (1963) states that, "One of the key ideas of Bayesian statistics is that probability is orderly opinion, and that inference from data is nothing other than the revision of such opinion in the light of relevant new information".

3.11.1 Bayesian Estimation

Bayesian estimation starts by making an assumption that the parameter space (ϕ) follows a probability distribution, and is considered to be random. The samples collected are then used to update the probability distribution or function.

Assuming the parameter space involves only one parameter (θ), then the function of the component (θ) on ϕ is referred to as the prior function which shows the experimenter's belief on (θ).

The conditional probability of the data based on the prior is termed the likelihood. The posterior function transforms to be the conditional function of the parameter after the sample data has been realized which sums up the prior and the sample information.

Taking Y as a dummy variable whose function relies on the coefficient (θ). Then the probability of Y is given as $P(y|\theta)$, Let the prior function be $\pi(\theta)$. The Posterior function of θ is obtained as follows:

$$P(\theta|y) = \frac{P(\theta,y)}{P_y(y)} = \frac{P(y|\theta)\pi(\theta)}{P_y(y)}$$

Since $P_y(y)$ does not depend on θ, we can show the posterior function as proportional ($\propto$) to (likelihood x prior function). That is:

$$P(\theta|y) \propto P(y|\theta)\pi(\theta) \tag{3.23}$$

Assuming $\vec{Y}$ is not a fixed vector from the function $P(y|\theta)$, then the posterior function is given as:

$$P(\theta|y) = \frac{P(y_1,y_2,...,y_n,\theta)\pi(\theta)}{P(y_1,y_2,...,y_n)}$$

$$= \frac{L(y_1,y_2,...,y_n|\theta)\pi(\theta)}{P(y_1,y_2,...,y_n)}$$

$$\propto L(y_1, y_2, ..., y_n|\theta)\pi(\theta) \tag{3.24}$$

where, $L(y_1, y_2, ..., y_n|\theta)$ is the combined form of the probability function of $\vec{Y}$ called the likelihood function.

When it comes to the Bayesian scenario, every information on θ realized in observed data and the prior information is held in the posterior function. Hence, the posterior function gives us estimates of θ that we can depend on compared to the prior. So, the Bayesian estimate of a coefficient is the posterior mean. This information can be applied in knowing the Bayesian estimates for several parameters.

From the equation below,

$$X_i = \beta_0 + \beta_1 Y_{i-1} + \beta_2 Y_{i-2} + \beta_3 Y_{i-3} + ... + \beta_p Y_{i-p} + \epsilon_i$$

$$= \vec{\beta}'\vec{Y_i} + \epsilon_i \tag{3.25}$$

where, $\vec{\beta}' = (\beta_1, \beta_2, \beta_3, ..., \beta_p)$ and $\vec{Y_{i-t}} = (Y_{i-1}, Y_{i-2}, Y_{i-3}, ..., Y_{i-p})'$

The $X_i's$ are independent with each having a normal distribution with mean $\vec{\beta}'Y_i$ and variance σ^2. So that the conditional distribution of the $X_i's$ given $\vec{\beta}$ is:

$$f(X_i|\vec{\beta}) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(X_i - \vec{\beta}'Y_i)^2} \tag{3.26}$$

; $|X_i| \geq 0$

The likelihood function is given as:

$$L(X|\vec{\beta}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(X_i - \vec{\beta}'\vec{Y})^2}$$

$$= (\frac{1}{\sqrt{2\pi\sigma^2}})^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}(X_i - \vec{\beta}'\vec{Y_i})^2} \tag{3.27}$$

Now say the prior function of the random vector $\vec{\beta}$ is the **multivariate normal distribution** with mean vector $\vec{\mu}$ and covariance matrix, $\sum$. Meaning $\vec{\beta} \propto N_p(\vec{\mu}, \sum)$. Then,

$$\pi(\vec{\beta}) = \frac{1}{|\sum|^{\frac{1}{2}}} (\frac{1}{\sqrt{2\pi}})^p e^{-\frac{1}{2}[Q(\beta)])} \tag{3.28}$$

where,

$$Q(\beta) = (\vec{\beta} - \vec{\mu})' \sum^{-1} (\vec{\beta} - \vec{\mu})$$

.Also,

$$\vec{\beta} = (\beta_1, \beta_2, \beta_3..., \beta_p) \quad , \quad \vec{\mu} = (\mu_1, \mu_2, \mu_3, ..., \mu_p) \quad \text{and} \quad \sum^{-1} = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \cdots & \vdots \\ a_{p1} & \cdots & a_{pp} \end{pmatrix}$$

Hence the posterior function is:

$$f(\vec{\beta}|X) \quad \propto \quad L(X|\vec{\beta}) \;\; y \;\; \pi(\vec{\beta})$$

$$\propto (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\vec{\beta}'Y_i)^2} \quad y \quad (2\pi)^{-\frac{p}{2}}|\sum|^{-\frac{1}{2}} e^{-\frac{1}{2}[(\vec{\beta}-\vec{\mu})'\sum^{-1}(\vec{\beta}-\vec{\mu})]}$$

$$= \quad K \;\; exp\, -\frac{1}{2} \;\; [\;\; \frac{1}{\sigma^2} \;\; \sum_{i=1}^{m}(X_i - \vec{\beta}'\vec{Y_i})^2 \quad + \quad (\vec{\beta}-\vec{\mu})' \sum^{-1}(\vec{\beta}-\vec{\mu}) \;\;] \qquad 3.29$$

K stands for all components that does not include $\vec{\beta}$

Now say,

$$\text{V}=\frac{1}{\sigma^2}\sum_{i=1}^{m}(X_i - \vec{\beta}'\vec{Y_i})^2$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{m}(X_i - \sum_{j=1}^{q}\beta_j Y_{ji})^2 \qquad 3.30$$

Let, i=1 and j=1,2,3 then :

$$(X_i - \sum_{j=1}^{3}\beta_j Y_{ji})^2 = (X_1 - \beta_1 Y_{11} - \beta_2 Y_{21} - \beta_3 Y_{31})^2$$

$$= \quad X_1^2 - 2X_1\beta_1 Y_{11} - 2X_1\beta_2 Y_{21} - 2X_1\beta_3 Y_{31} + \beta_1^2 Y_{11}^2 + 2\beta_1\beta_2 Y_{11}Y_{21}$$

$$+2\beta_1\beta_3 Y_{11}Y_{31} + \beta_2^2 Y_{21}^2 + 2\beta_2\beta_3 Y_{21}Y_{31} + \beta_3^2 Y_{31}^2 \qquad 3.31$$

From the above equation (3.30) can be summarized as follows:

$$V = \frac{1}{\sigma^2}\sum_{i=1}^{m} X_1^2 + \frac{1}{\sigma^2}\sum_{i=1}^{m}\sum_{j=1}^{q}\beta_j^2 Y_{ji}^2 - 2\frac{1}{\sigma^2}\sum_{i=1}^{m}\sum_{j=1}^{q} X_i\beta_j Y_{ji} + 2\frac{1}{\sigma^2}\sum_{i=1}^{m}\sum_{j=1}^{q-1}\sum_{s\neq j}^{q}\beta_j\beta_s Y_{ji}Y_{si} \qquad 3.32$$

Also let:

$$T = (\vec{\beta}-\vec{\mu})'\sum^{-1}(\vec{\beta}-\vec{\mu}) \qquad 3.33$$

Now let

$$\vec{\beta} = (\beta_1, \beta_2, \beta_3) \;\; , \quad \vec{\mu} = (\mu_1, \mu_2, \mu_3) \quad \text{and} \quad \sum^{-1} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Then equation (3.33) can be expressed in the following way:

$$(\beta_1-\mu_1 \ , \quad \beta_2-\mu_2, \quad \beta_3-\mu_3)\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} \beta_1-\mu_1 \\ \beta_2-\mu_2 \\ \beta_3-\mu_3 \end{pmatrix}$$

$$= \ a_{11}(\beta_1-\mu_1)^2+a_{21}(\beta_2-\mu_2)(\beta_1-\mu_1)+a_{31}(\beta_3-\mu_3)(\beta_1-\mu_1)$$

$$+a_{12}(\beta_1-\mu_1)(\beta_2-\mu_2)+a_{22}(\beta_2-\mu_2)^2+a_{32}(\beta_3-\mu_3)(\beta_2-\mu_2)+a_{13}(\beta_1-\mu_1)(\beta_3-\mu_3)$$

$$+a_{23}(\beta_2-\mu_2)(\beta_3-\mu_3)+a_{33}(\beta_3-\mu_3)+a_{33}(\beta_3-\mu_3)^2 \qquad 3.34$$

With reference from equation (3.34), T (ie. equation 3.33) can be expressed as:

$$T=\sum_{j=1}^{q}a_{jj}(\beta_j-\mu_j)^2+2\sum_{j=1}^{q-1}\sum_{s\neq j}^{q}a_{js}(\beta_j-\mu_j)(\beta_s-\beta_s)$$

$$=\sum_{j=1}^{q}a_{jj}(\beta_j^2+\mu_j^2-2\beta_j\mu_j)+2\sum_{j=1}^{q-1}\sum_{s\neq t}^{q}(a_{js}\beta_j\beta_s-a_{js}\beta_j\mu_s-a_{js}\mu_j\beta_s+a_{js}\mu_j\mu_s)$$

$$=\sum_{j=1}^{q}a_{jj}\beta_j^2+\sum_{j=1}^{q}a_{jj}\mu_j^2-2\sum_{j=1}^{q}a_{jj}\beta_j\mu_j+2\sum_{j=1}^{q-1}\sum_{s\neq j}^{q}a_{js}\beta_j\beta_s$$

$$-2\sum_{j=1}^{q-1}\sum_{s\neq j}^{q}a_{js}\beta_j\mu_s-2\sum_{j=1}^{q-1}\sum_{s\neq j}^{q}a_{js}\mu_j\beta_s+2\sum_{j=1}^{q-1}\sum_{s\neq j}^{q}a_{js}\mu_j\mu_s \qquad 3.35$$

Now:

$$\text{V+T}=\sum_{j=1}^{q}[\tfrac{1}{\sigma^2}\sum_{i=1}^{m}Y_{ji}^2+a_{jj}]\beta_j^2+2\sum_{j=1}^{q-1}\sum_{s\neq j}^{q}[\tfrac{1}{\sigma^2}\sum_{i=1}^{m}Y_{ji}Y_{si}+a_{js}]\beta_j\beta_s$$

$$-2\sum_{j=1}^{q}[\frac{1}{\sigma^2}\sum_{i=1}^{m}Y_{ji}X_i+m]\beta_j+R \qquad 3.36$$

where, $m = \sum_{j=1}^{q} a_{jk}\mu_k$, j=1,2,3,...,q and R is a constant term independent of β_j. Clearly,

$$\pi(\beta_{\sim}|X) = Ke^{-\frac{1}{2}Q_{\beta}(\beta)} \tag{3.37}$$

Therefore the Posterior function is a **Multivariate Normal Distribution**. Assuming the matrix definition of $Q_{\beta}(\beta)$ is $\sum_{\beta}^{-1}$, whereby $\sum_{\beta}$ is a P x P matrix which has an inverse $\sum_{\beta}^{-1}$.

The components of $\sum_{\beta}^{-1}$ is:

$$n_{jj} = \frac{1}{\sigma^2}\sum_{i=1}^{m} Y_{ji}^2 + a_{jj}, j = 1, 2, ..., q \tag{3.38}$$

$$n_{js} = \frac{1}{\sigma^2}\sum_{i=1}^{m} Y_{ji}Y_{si} + a_{js}, j \neq s \tag{3.39}$$

The coefficient component is expressed as a column vector $C_{\sim} = (c_1, c_2, ..., c_q)'$ of order q with the j^{th} element given as:

$$c_j = -2[\frac{1}{\sigma^2}\sum_{i=1}^{m} X_i Y_{ji} + \sum_{k=1}^{q} a_{jk}\mu_k], j = 1, 2, ..., q \tag{3.40}$$

The mean vector associated with the Posterior function is:

$$\vec{\mu_{\beta}} = -\frac{1}{2}\sum_{\beta}\vec{C} \tag{3.41}$$

Even though the estimates of the Bayesian model coefficients in (3.34) are given as:

$$\hat{\vec{\beta}} = \vec{\mu_{\beta}} \tag{3.42}$$

3.12 Methods of Analysis

VARs or BVARs are mostly used to monitor changes in data. It is challenging to explain in simple terms. It is tough to make inference on them by examining the parameters in the regression model. The estimated parameters on subsequent lags seems to swing and the presence of large cross-equation results. Due to this, the best method available would be Impulse Response Funtions (IRFs) and Forecast Error Variance Decomposition (FEVD).

3.12.1 Impulse Response Function (IRF)

Impulse Response Functions identify the outcome of a one standard deviation turbulence to one of the VAR residuals on recent and projected values of the other variables given that this residual approach zero in ensuing period whilst all other residuals are zero. Take a bivariate VAR in reduced form as

$$y_t = \beta_0^{-1} A y_{t-1} + \beta_0^{-1} \mu_t \ , \ \beta_0^{-1} \mu_t = e_t$$

. y_t is a vector made up of x_t and z_t

The IRF to a unit shock in μ_t can be approximated once we know β_o^{-1}.Assume we want to identify the outcome of the first variable in the system during period=1,2,... when a turbulence hits the system at time 0, then

$$\mu_o = \begin{bmatrix} 1 \\ \\ 0 \end{bmatrix} \text{ and } y_o = \begin{bmatrix} x_o \\ \\ z_o \end{bmatrix} = \beta_o^{-1} \mu_o$$

For all>0, $y_s = \beta_0^{-1} A y_{s-1}$

Thus IRF is a down to earth way of representing the behaviour of variables in a VAR to various shocks in the residual terms overtime. They are the time paths of the variables in response to shocks in the system. IRFs are functions of the

approximated VAR parameters. In a K-variable system, there are K^2 impulse response functions.

3.12.2 Forecast Error Variance Decomposition (FEVD)

It allocates the variance of a variable into parts that are divided to each of the set of shocks. It also elaborate the quantity of information each variable contributes to the other variables in the VAR. It gives the relation of the role of various variables in explaining the changes in a certain variable in a system.

3.12.3 Bayesian Credible Interval

The estimates of the parameters and forecasts of the Bayesian model are interpreted as follows;

The set $C \subseteq \Omega$ is a $100(1-\alpha)\%$ credible set with respect to $P(\omega|y)$ if:

$$P(\omega \in C|y) = \int_c P(\omega|y)dw = 1 - \alpha \tag{3.43}$$

For example, suppose $\omega = g(\beta) = \beta_j$, being a single coefficient.Then a 95% credible interval for β_j is any interval, [a,b] such that;

$$P(a \leq \beta_j \leq b|y) = \int_a^b P(\beta_j|y)d\beta_j = 0.95 \tag{3.44}$$

3.13 MCMC Methods

It is a type of code that is used in sorting from a probability function dependent on creating a Markov chain that contains the right function as its equivalent function.

The condition of the Markov chain after making some movements is now taken as a portion of the right distribution.

The samples or selection become refined as a mechanism of the number of movements. Again, the samples of the integrand made by chance from the MCMC method are same so the Markov Chain is conducted in a special form to have the integrand as its function.

Steps of an MCMC algorithm

Algorithm 1 MCMC alogorithm

-
- Step 1: Run a typical long : $(\phi^1, \phi^2, ..., \phi^i)$ (which follows a continous replication)
-
- step 2: Estimate the extra amount added $E(h(\phi|y)) = \sum_{\phi} h(\phi)\pi(\phi|y)$
  By the arithmetic averages $h_m = \frac{1}{L}\sum h(\phi^i))$
-
- step 3: Calculate standard error $\bar{h_m}$ of the Monte Carlo to setup an asymptotically correct confidence for $E(h(\phi|y))$ There are lots of MCMC methods, examples are the Random Walk Monte Carlo. Examples of the Random Walk Monte Carlo methods are: Gibbs sampling, Slice sampling, etc.

=0
=0

3.14 GIBBS Sampling

It is an MCMC code for accessing a series of responses estimated from a particular multivariate probability function, where its hard to do a direct sampling.

This series may be applied to estimate the joint distribution. Example to calculate the marginal function (distribution) of one of the variable or a sample of the variables.

Relating to other MCMC algorithms,the Gibbs sampler sets up a Markov series of subsets,where each matches with close subsets. Hence, it has to be done carefully if separate samples are needed.

Gibbs sampling best fits selecting the posterior function of a Bayesian mapping or network, since Bayesian mappings are distinctively identified as a set of conditional distribution as demonstrated in (Anno-Kwakye, 2017).

3.14.1 Gibbs Sampling Framework

The idea of Gibbs selection is provided a multivariate function, it is much easier to select from a conditional distribution than to marginalize by integrating over a total distribution. Say, we need to get z subsets of $Y = (y_1, ..., y_n)$ from a joint distribution $p(y_1, ..., y_n)$.
Let the i^{th} subset be $Y^{(i)} = (y_1^{(i)}, ..., y_n^{(i)})$. We continue in the following way:

1. Start with a value $y^{(i)}$

2. We want to proceed with the next selection.Say this next subset is $y^{(i+1)}$.

   Since $y^{(i+1)} = (y_1^{(i+1)}, y_2^{(i+1)}, ..., y_n^{(i+1)})$ is a vector, we select every aspect of the vector, $y_j^{(i+1)}$, from the function of that component based on

the rest of the parts selected so far. But there is a hindrance which is answered by basing on $y^{(i+1)}$'s parts till $y_{j-1}^{(i+1)}$'s parts and continue basing on $y^{(i)}$ parts beginning from $y_{(j+1)}^{(i)}$ to $y_n^{(i)}$. In order to arrive at that, we select the parts in order, beginning from the first part. Officially,selecting $y_j^{(i+1)}$we keep abreast in relation to the function pointed out by $p(y_j^{(i+1)}|y_1^{(i+1)}, ..., y_{j-1}^{(i+1)}, y_{(j+1)}^{(i)}, ..., y_n^{(i)})$.

We do not gain if the value that the $(j+1)^{th}$ part had in the (i^{th}) selection not the $(i+1)^{th}$ selection is applied.

3. Redo the previous procedure k times. If this type of selection is conducted, these facts are valid:

   - The selections made estimates the combined function of all variables.
   - The marginal function of any sample of variables can be estimated by easily accommodating the selections for that samples of variables, forgoing the rest.
   - The projected response of any variable can be estimated by averaging over all the selections

The estimates of the errors can be measured numerically using a Central Limit Theorem (CLT).In particular, we obtain ;

$$\sqrt{S}(\hat{g}_s - E[g(\beta)|y]) \rightarrow N(0, \sigma_g^2) \tag{3.45}$$

As S$\rightarrow \infty$, $where\ \ \sigma_g^2 = var(g(\beta)|y)$. Using this estimate in equation (3.43) and the properties of the normal density,we can write

$$P(E[g(\beta)|y] - 1.96\frac{\hat{\sigma}_g}{\sqrt{S}} \leq \hat{g}_s \leq E[g(\beta)|y] + 1.96\frac{\hat{\sigma}_g}{\sqrt{S}}) \approx 0.95 \tag{3.46}$$

We can then rearrange the probability statement in equation (3.44) to find an approximate 95% confidence interval for $E[g(\beta)|y]$ of the form

$$[\hat{g}_s - 1.96\frac{\hat{\sigma}_g}{\sqrt{S}}, \hat{g}_s + 1.96\frac{\hat{\sigma}_s}{\sqrt{S}}]$$

3.15 Conclusion

This chapter mainly emphasized on the analytical procedures taken to draw inferences from the available data. The origin of the data and manipulations made on the data to make it suitable for quantitative analysis was explained.

The ARIMA, EGARCH and BVAR models were specified as the suitable, alternate or complement models for the statistical analysis by the researcher. The ARIMA model was the first been used on the data since the model has components that can account for critical indicators in the data such as differencing, stationarity, lag order, and seasonality which are of interest to the researcher.

In the case where there is an ARCH effect or heteroscedasticity present in the data, the ARIMA model fails to account properly, otherwise it good for modelling and forecasting data existing in linear form.
If ARCH effect is established, EGARCH is used as an alternative model for capturing the volatilities in the residuals of the data. One key property is its ability to detect the assymmetries in the data and forecast the conditional variance as well. In the event where it is not able to do the forecasting appropriately, the BVAR is used for the forecasting aspect.

The BVAR is perceived to have the ability to forecast appropriately by the researcher hence is used. It serves as a multi-purpose model by studying the relationship of multiple variables and predicting as well. BVAR models are extension of VAR models which functions in the same way as VARs but BVARs are improvements of the VARs with the assumption that the parameters have a probability distribution.

Moreover, VARs are over-parameterized which violates the parsimonious rule in modelling hence not of much interest to the researcher. Relevant statistical test were used to assess each model and methods; such as the MLE was used to estimate the parameters of the EGARCH, while the MCMC methods were used to simulate and estimate the BVAR model.

The best models were selected based on criteria such as AIC, BIC, and H-Q, and assessed using loss functions such as RMSE and MAE.

Chapter 4

DATA ANALYSIS AND DISCUSSIONS

4.1 Introduction

This Chapter focused on the analysis of the data. Practical aspects of the theories discussed in Chapter three are implemented here. Also, the outcomes of the analysis and the data processing tools employed are presented here.

4.2 Results

4.2.1 Descriptives

Table 4.1 shows the quantitative explanations of the exchange rate data under observation. The mean gives us the average value recorded from the data, while the standard deviation demonstrates the level of dispersion of the observations from the average.

The data consists of monthly exchange rates of the cedi against the US dollar over a period of 14 years. It was obtained from the Central Bank's website. The data spanned from January, 2008 to March, 2022 with 171 observations in total.

It was then split into training and testing sets. The training set composed of 80% of the data and was used in the estimation of the model parameters, while the testing set which composed of 20% of the data was used in the projection of future exchange rates. The training set which spanned from January, 2008 to December, 2018 was made up of 136 observations, while the testing set which

was from January, 2019 to March, 2022 consisted of 35 observations.

The plot of the data showed a continuous upward trajectory of the rate of exchange between the cedi and the dollar which obviously showed no stationarity as shown in Figure 4.1.

Table 4.1: Description of the data

Statistic	Minimum	Maximum	Mean	Standard Deviation	skewness
Exchange rate	0.97	7.11	3.31	1.70	0.18
Statistic	kurtosis	Jarque-Bera	P-value	Observations	
Exchange rate	-1.41	14.986	0.0006	171	

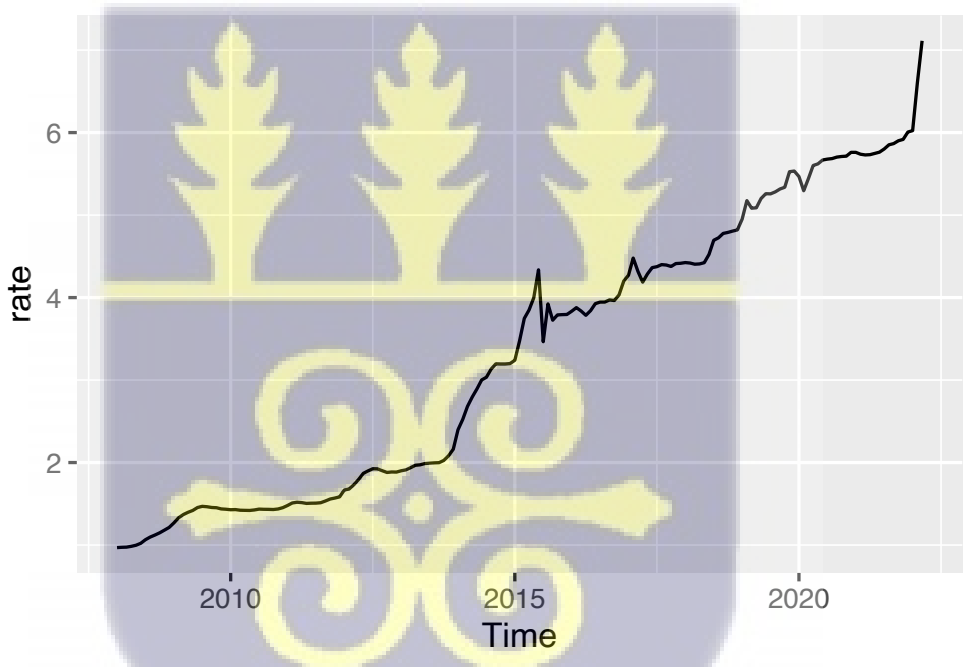


Figure 4.1: **Plot of exchange rate of the Cedi to the US dollar over time**

One desirable property in time series is to make the data level stationary. From Figure 4.2, it was obvious that the plot showed a trend with spikes that was consistent over the time period, meaning it became stationary after being differenced once. Moreover, the volatility clustering property of time series is evident. This study focused on the returns of the exchange rate data as indicated in equation (3.1) of Chapter 3.

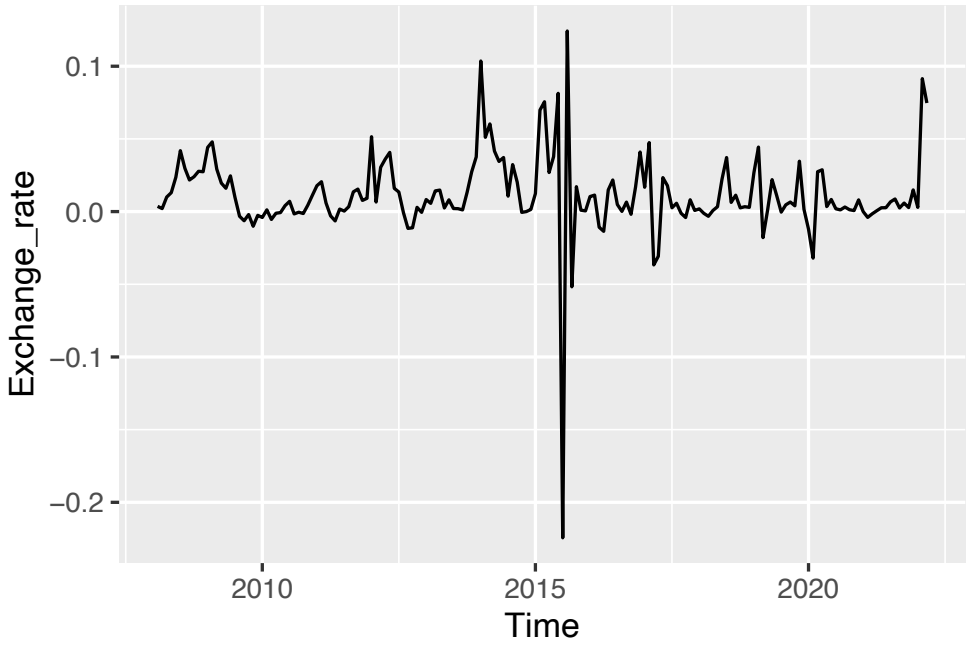


Figure 4.2: **Plot of the exchange rate data after first differencing**

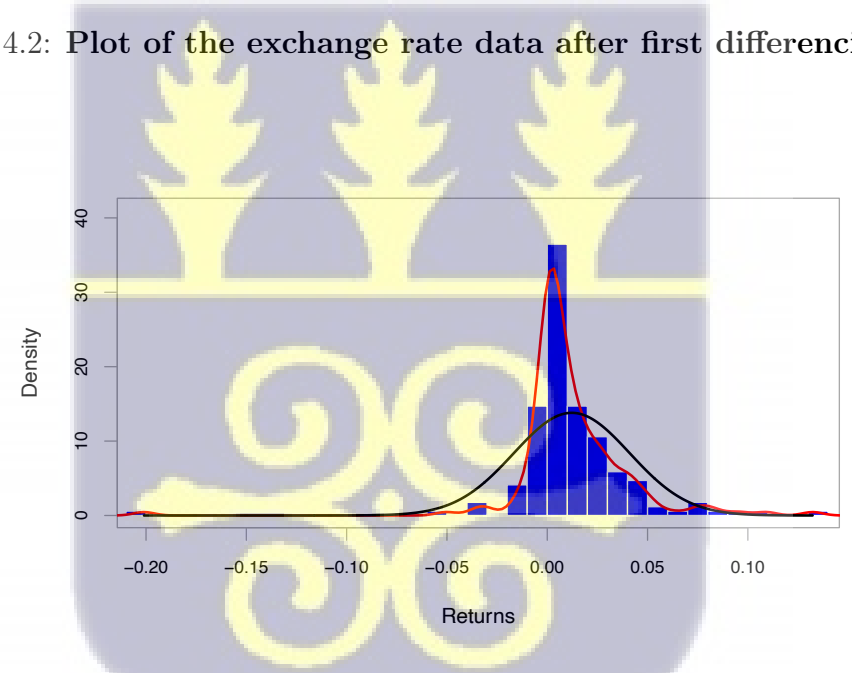


Figure 4.3: **Histogram of the exchange rates after first differencing**

4.2.2 Unit Root Test of the Data

To confirm stationarity of the time series data, the Augmented Dickey-Fuller and KPSS test were performed. The results of the two tests are outlined in Tables 4.2 and 4.3 respectively. From Table 4.2, the null hypothesis of non-stationarity or presence of unit root is rejected. Otherwise in Table 4.3, we fail to reject the null

hypothesis of level stationary of the KPSS test. This was because the p-value of 0.01 (2 d.p) was below the 5% significance level for the ADF test, while that of the KPSS test was 0.1 supporting the null hypotheses of level stationary. This translates to mean that the data is stationary.

Table 4.2: Unit root test of the returns

Statistic	Test statistic	P-value
Augmented Dickey-Fuller	-3.955	0.013

Table 4.3: Unit root test of the returns

Statistic	Test statistic	P-value
KPSS	0.1186	0.1

Table 4.4: Description of the exchange rate data

Statistic	Minimum	Maximum	Mean	Standard Deviation	skewness
Exchange rate	-0.200	0.132	0.012	0.006	-1.300
Statistic	kurtosis	Jarque-Bera	P-value	Observations	
Exchange rate	18.425	2520.4	0.0006	171	

From Table 4.4, the kurtosis value is positive meaning a "heavy-tail". The skewness shows whether there is symmetry in the data or not.

Usually, Normal distributions have skewness value of zero. A positive value for the skewness means the data is more skewed to the right. The value obtained from the analysis was negative, which shows that the data is more skewed to the left.

Jarque-Bera statistic was 2520.4 with a p-value less than 0.05 which suggests that we reject the null hypotheses of normality.

4.2.3 Q-Q Plot of the Data

Once it is evident that there is serial correlation, we continue to check the distribution of the residuals of the data. For us to know what distribution the residuals follow, the plot of the quantile (Q-Q) is applied. This plot shows time series together with the best fitted distribution where possible. Figure 4.4 depicts a slight deviation from the standard normal distribution. The quantile plot exhibited a "heavy-tail" at the far right end of the straight line.

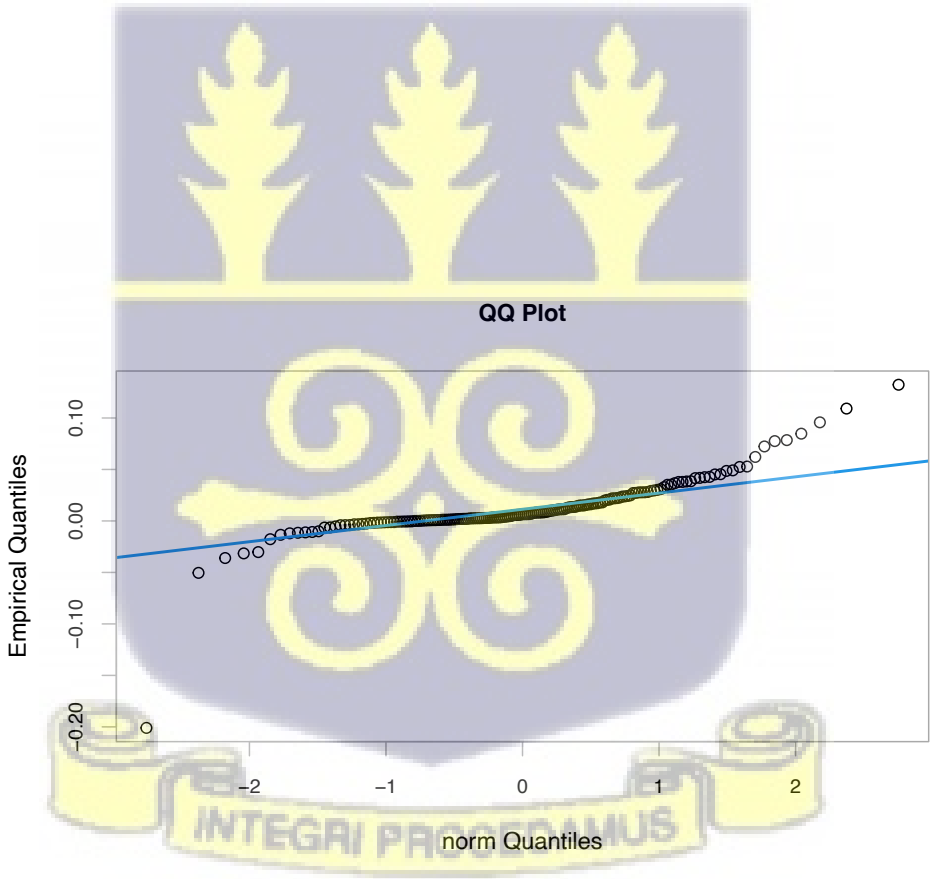


Figure 4.4: **Quantile plot of exchange rate data after being differenced**

4.3 Model Selection

In order to forecast volatility of the exchange rates, we have to select a model with an optimal lag length. The information criterion aids in this process, that is, AIC, BIC and AICc which were explained in the previous Chapter.

All we need is a model with the least AIC, AICc and BIC which then becomes the best model to fit our data. The ACF and PACF plot helps to accommodate a quantity of successive pairs of the ARMA(p,q). All the models had a d-value of 1 because our returns series was differenced once, and the values we considered for p and q were because the plot showed those lags at significant points.

We chose the model with least value of AIC as the best model. Thus $ARIMA\ (0,1,1)$ is considered as the best for the projection of future exchange rates.

To determine how effective our chosen model is, a diagnosis of the model is carried out which will further help us to know whether to move on with the ARCH family models.

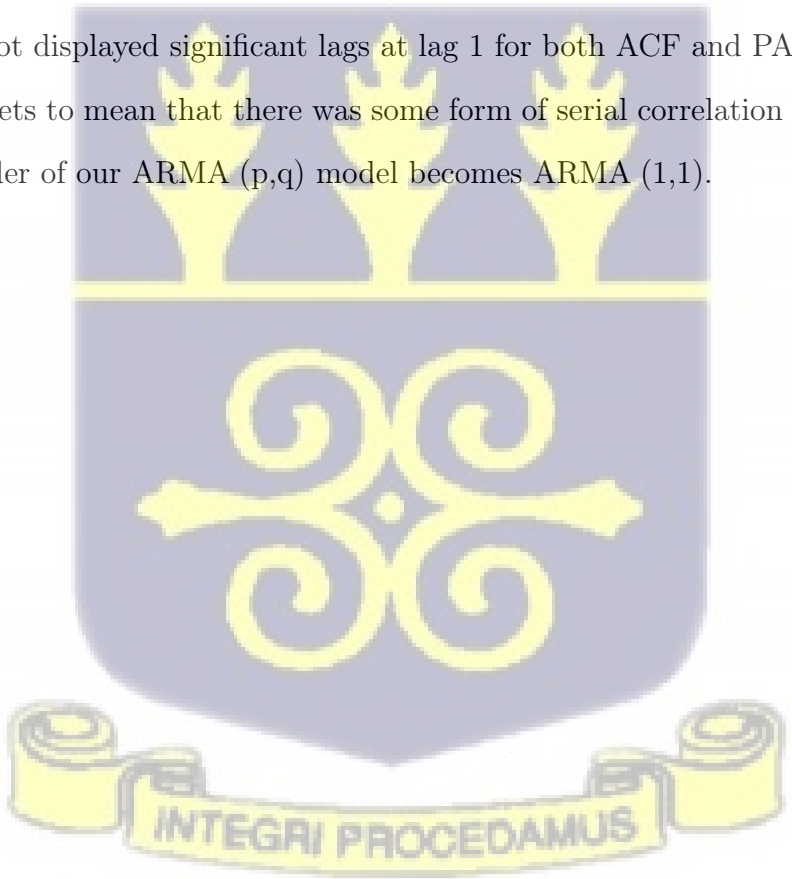
Table 4.5: ARIMA model selection

Model	AIC
(0,1,0)	272.965
(0,1,0)	275.167
(0,1,1)	258.361
(0,1,1)	260.474
(0,1,2)	262.894
(0,1,3)	262.438
(0,1,4)	263.719
(0,1,5)	265.668
(1,1,0)	262.237
(1,1,1)	260.612
(1,1,2)	263.019
(2,1,0)	261.393
(2,1,0)	263.789
(2,1,1)	262.290
(3,1,0)	263.829
(3,1,0)	266.357
(4,1,0)	264.796
(4,1,0)	267.403
(4,1,1)	266.285
(5,1,0)	267.059
(5,1,0)	269.769

4.3.1 Autocorrelation for the Exchange Rate Data

To confirm the autocorrelation or serial correlation which forms the basis for the execution of the ARCH or GARCH models, we use the plots of the ACF and PACF. It also helps us determine the order of our ARMA(p,q) model. Our data has to exhibit some form of correlation before the ARCH or GARCH model can be undertaken.

From the plot of Figure 4.5 and 4.6, a bar that goes pass the significant limit means that there exist some form of serial correlation at that specific lag. Otherwise, if all lags are clustered around zero and the Ljung-Box Q statistic at all lags are not significant, we conclude that the data shows no serial correlation. Our plot displayed significant lags at lag 1 for both ACF and PACF plots which interprets to mean that there was some form of serial correlation in the data and the order of our ARMA (p,q) model becomes ARMA (1,1).



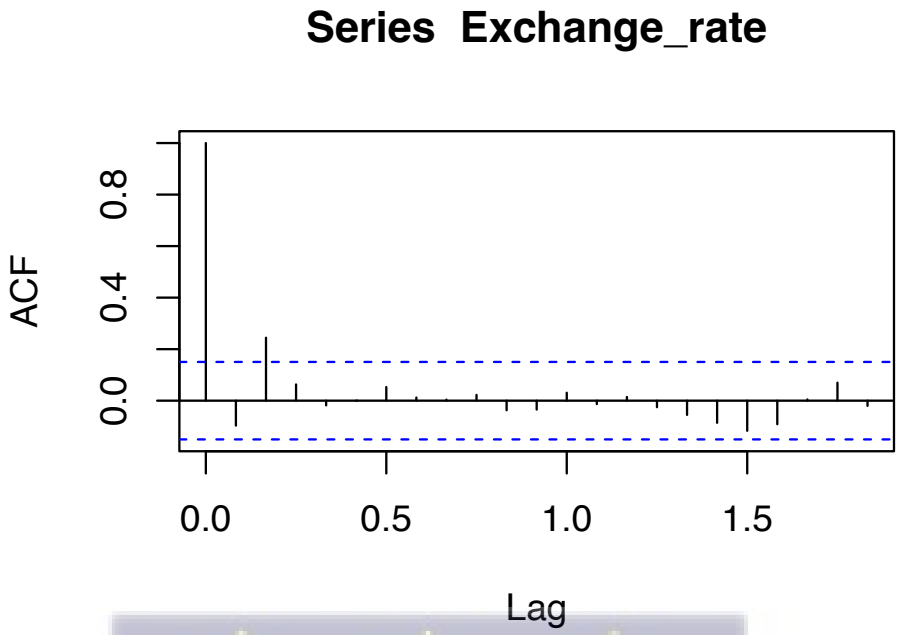


Figure 4.5: **ACF plot of exchange rate**

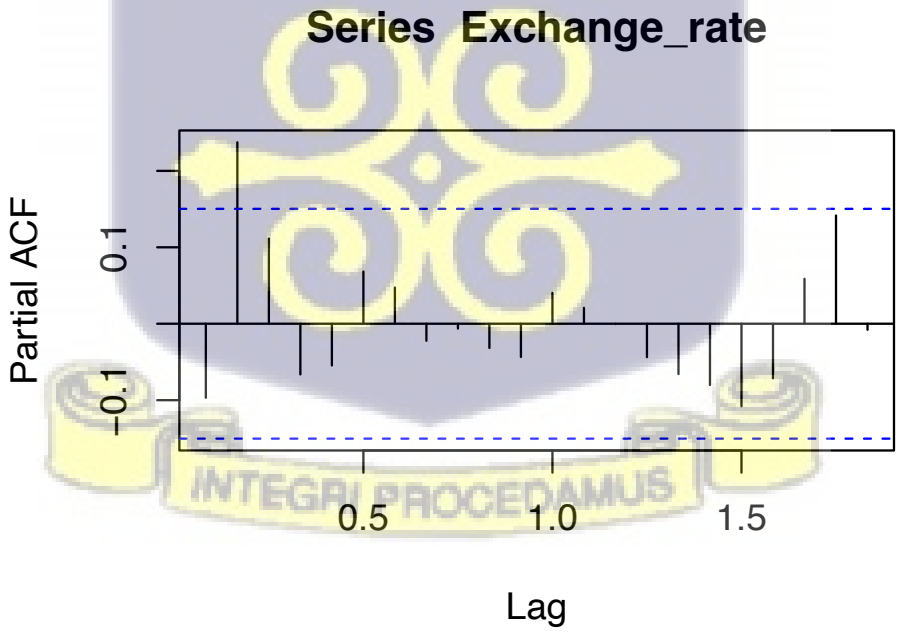


Figure 4.6: **PACF plot of exchange rate**

As it known that ARIMA models are good at modelling data that exists in a linear form and do not consider any new signals, all they do is to give the outstanding linear forecast. However, they do not function well in the prediction of non-linear data as shown in Figure 4.2. That is in order for us to fit the volatility well, the suitable models are the ARCH family models.

Whether to carry on or not with the ARCH family models, a test is needed to give us a first hand information. Hence, the ARCH-LM statistical test is used which has a null hypotheses of no ARCH effect in the residuals of our ARIMA model against the hypotheses of the presence of ARCH effect. The Ljung-Box test is carried out as well.

Table 4.6: Test for autocorrelation of $ARIMA\ (0, 1, 1)$

Statistic	Test Statistic	P-value
Ljung-Box Test	14.282	0.027

Observing Table 4.6, the test statistic recorded had a corresponding p-value of 0.027 which was below the 5% significance level, suggesting that we reject the null hypotheses of no serial correlation. Hence, there was autocorrelation in the series.

Table 4.7: Test for ARCH effect of $ARIMA\ (0, 1, 1)$

Statistic	Test Statistic	P-value
ARCH-LM Test	42.887	0.016

Again, Table 4.7 which showed the outcome of the ARCH-LM test suggested that the null hypothesis should be rejected since the p-value of 0.016 was below the significance level. Hence, the ARCH effect was present in our series.

Table 4.9 gives us the estimates of the parameters of the ARIMA model with mean equation of ARMA(1,1). The corresponding error measures were minute which suggested that our model parameters were significant.

Table 4.8: Estimates of $ARIMA\ (0, 1, 1)$

Variable	Estimate	Std. Error
MA(1)	-0.674	0.127
σ^2	21.76	

Table 4.9: Diagnostic criterion of $ARIMA\ (0, 1, 1)$

Statistic	Value
AIC	258.06
AICc	258.36
BIC	261.58

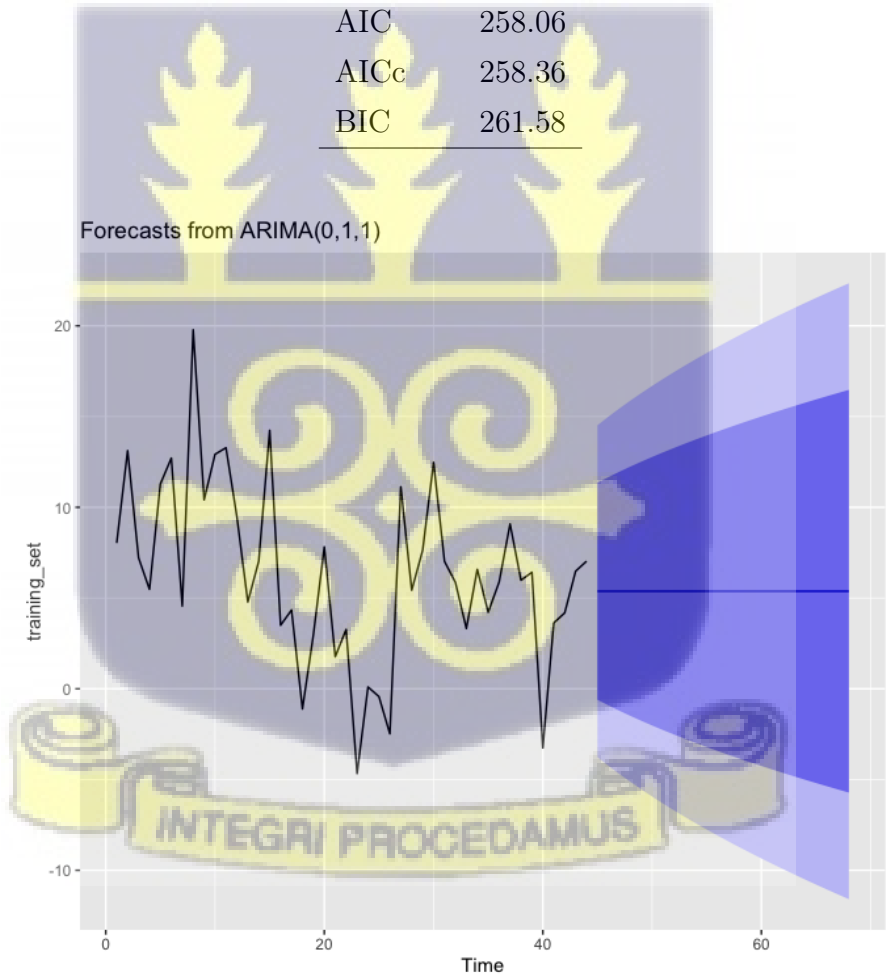


Figure 4.7: **Forecast plot of the ARIMA model**

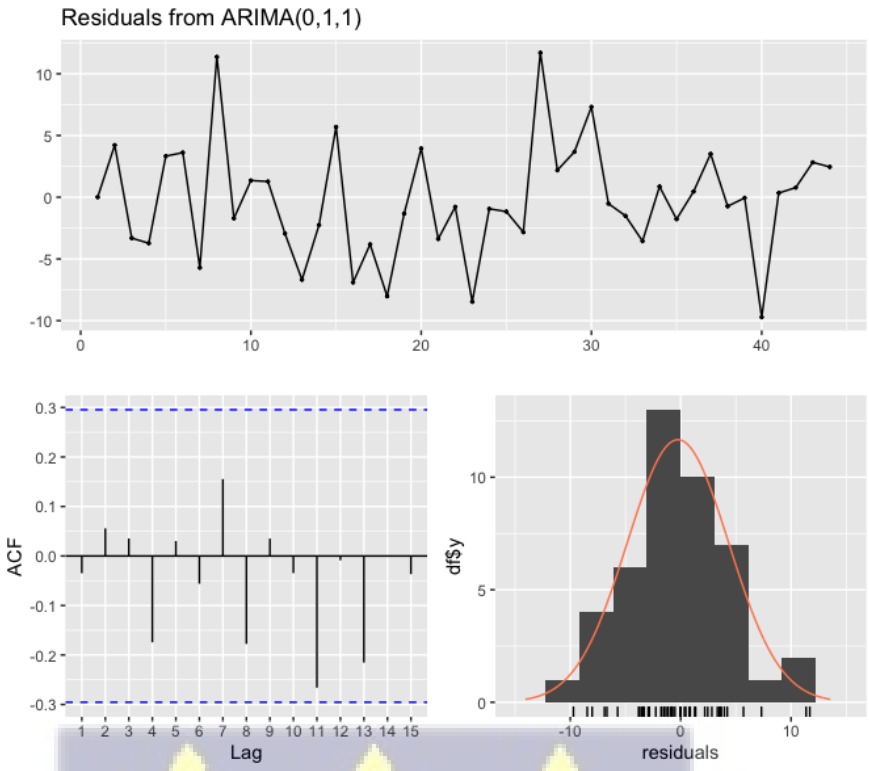


Figure 4.8: **Residual plot of the forecast from the ARIMA model**

Table 4.10: Error measures of $ARIMA\ (1,1,1)$

Statistic	Value
RMSE	4.557
MAE	3.472

The ARCH family models can now be fitted following the confirmation of the ARCH effect. Hence, this study was centered only on the EGARCH and standard GARCH type of the ARCH family models. Our proposed model (EGARCH) will be compared to only the standard GARCH as the standard GARCH is the widely used model in financial time series. The other types of the family models were not of concern.

4.3.2 A Comparison of the EGARCH Model to the Standard GARCH model

EGARCH (2,1) emerged as the best model to fit the variance in the residuals among the EGARCH models generated, as it is clear from Table 4.11 that EGARCH (2,1) had the lowest values for the AIC and BIC . This suggests that the model might produce some desirable outcomes from the data, which would lead to better estimates of our parameters in the model and subsequently better forecast of the variances in the residuals.

Table 4.11: EGARCH model selection based on their AIC,BIC and H-Q

Model	AIC	BIC
EGARCH(1,1)	-4.9639	-4.8354
EGARCH(1,2)	-4.9867	-4.8367
EGARCH(2,1)	-5.0361	-4.8648
EGARCH(2,2)	-5.0342	-4.8414

Assuming a normal approximation for the distribution of the residuals, the estimates of the parameters of the selected model are displayed in Table 4.12 below. It is obvious that the parameters of the model were significant as estimated with corresponding p-values less than the significance level of 0.05. This shows that the EGARCH (2,1) model would do well in describing the variances in the residuals and would give good outcomes when used.

Meanwhile, the estimates of the standard GARCH model from Table 4.14 were also significant and a measure of the error from Table 4.15 showed that the estimates of the model were near accuracy for the data at hand but EGARCH (2,1) was slightly better. This suggests the accuracy of the Standard GARCH were somehow close to that of EGARCH (2,1) but not as accurate as that of the EGARCH model.

Table 4.12: Estimates of EGARCH(2,1)

Variable	Estimate	Std. Error	test statistic	Prob.
μ	-0.002278	0.001557	-1.46314	0.143429
$ma1$	0.411497	0.082994	4.95813	0.000001
ω	-2.142669	0.556355	-3.85126	0.000118
Alpha1	-0.032534	0.148819	-0.21862	0.826948
Alpha2	0.540392	0.132636	4.07425	0.000046
Beta1	0.710892	0.069651	10.20642	0.000000
gamma1	0.885065	0.189407	4.672830	0.000003
gamma2	0.230853	0.199853	1.15511	0.248043

Table 4.13: Error measures of $EGARCH(2,1)$

Statistic	Value
RMSE	0.0326
MAE	0.0175

Table 4.14: Estimates of GARCH(1,1)

Variable	Estimate	Std. Error	test statistic	Prob.
μ	-0.005678	0.001797	-3.1597	0.001579
$ma1$	0.584658	0.060134	9.7226	0.000000
ω	0.000132	0.000037	3.5898	0.000331
Alpha1	0.933537	0.153996	6.0621	0.000000
Beta1	0.065463	0.041493	1.5777	0.114635

Table 4.15: Error measures of $GARCH(1,1)$

Statistic	Value
RMSE	0.0399
MAE	0.0811

Now that the best model (EGARCH (2,1)) has been identified and evaluated, we now have to pass this model through an assessment to ascertain the degree or

level to which this model suits the data (model diagnostics check). These checks are conducted on the residuals of EGARCH(2,1).

4.3.3 Diagnostic Checking of Best Performing Model

The following tests were carried out on the residuals of the model to ascertain if the model could be good in forecasting. We also verified if there was still the presence of serial or autocorrelation in the residuals. Table 4.17 proved that the null hypothesis of no serial correlation was not rejected, giving the Ljung-Box Q statistic of 0.226 and p-value of 0.635 which was above 5% level of significance. This gave some evidence that serial correlation was absent in the residuals of the model.

The ARCH-LM test from Table 4.16 also supported the fact that there was no ARCH effect present in the residuals, giving the p-value of 0.441 hence, the null hypothesis could not be rejected. There was now sufficient evidence that there were no heteroscedasticity among the residuals.

Table 4.16: Test for autocorrelation of EGARCH (2,1)

Statistic	test statistic	P-value
ARCH-LM Test	0.595	0.441

Table 4.17: Test for autocorrelation of EGARCH (2,1)

Statistic	test statistic	P-value
Ljung-Box Test	0.226	0.635

Figure 4.9 showed no significant lags for the squared standardized residuals after fitting the EGARCH model. The check for the normality assumption of the residuals was the last of our assessment.

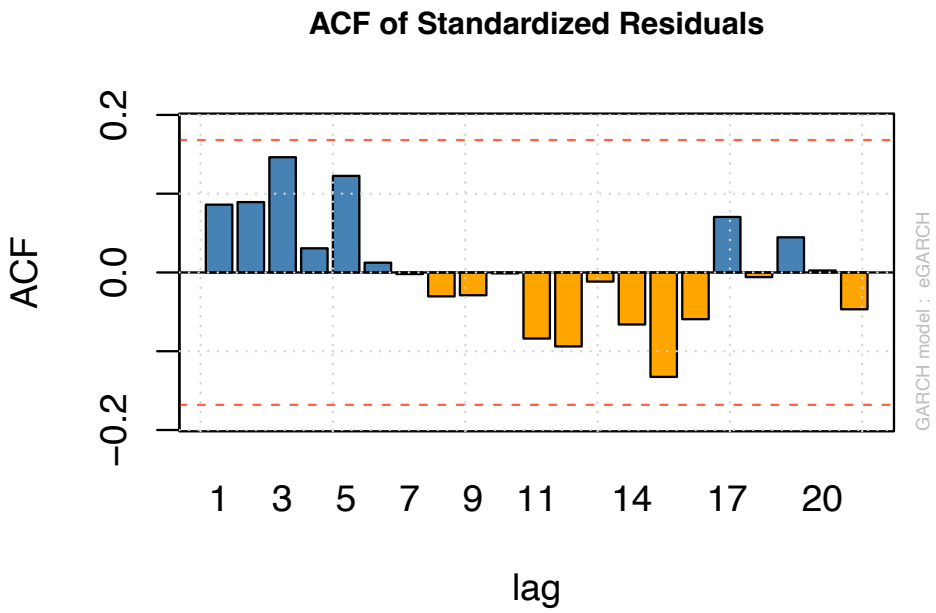


Figure 4.9: **Residual plot of the EGARCH(2,1) model**

The outcome is shown from the Q-Q plot of Figure 4.10. Most of the residuals lay on the straight line with little deviation at the tail end, signaling that the normality assumption holds for the model. The density plot also exhibited the characteristics of a normal distribution as shown in Figure 4.12.

Figure 4.11 on the other hand is the forecast plot of the EGARCH. We could see at the far right end of the plot that it produced a straight line with exact bounds (the yellow coloured region) demonstrating a regular linear path. This indication from the plot tells us that the EGARCH (2,1) model is a consistent estimator of our forecasted values which means the model possess one of the good properties of an estimator.

The model selected had passed all the assessment carried out hence would be good for fitting the variance in the residuals.

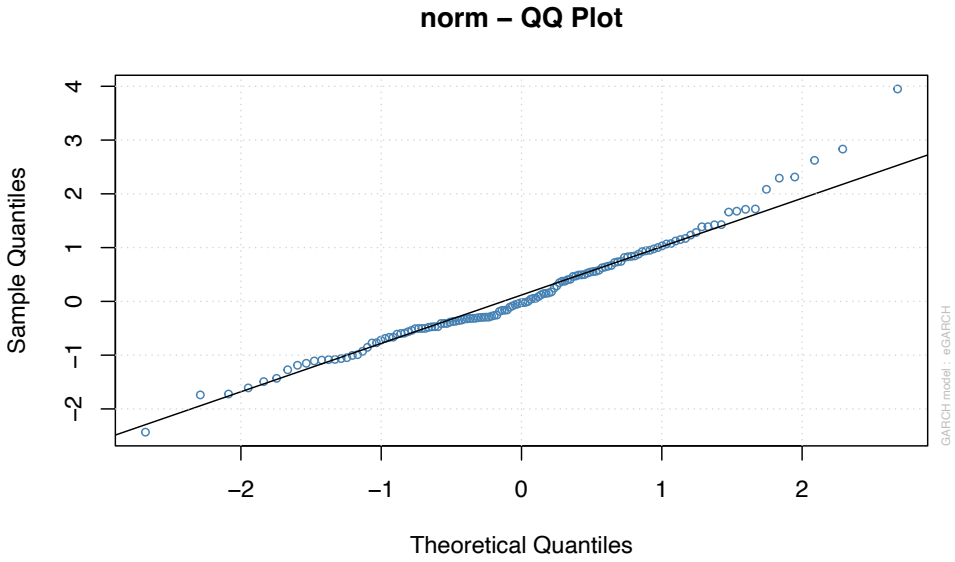


Figure 4.10: **Q-Q plot of the EGARCH(2,1) model**

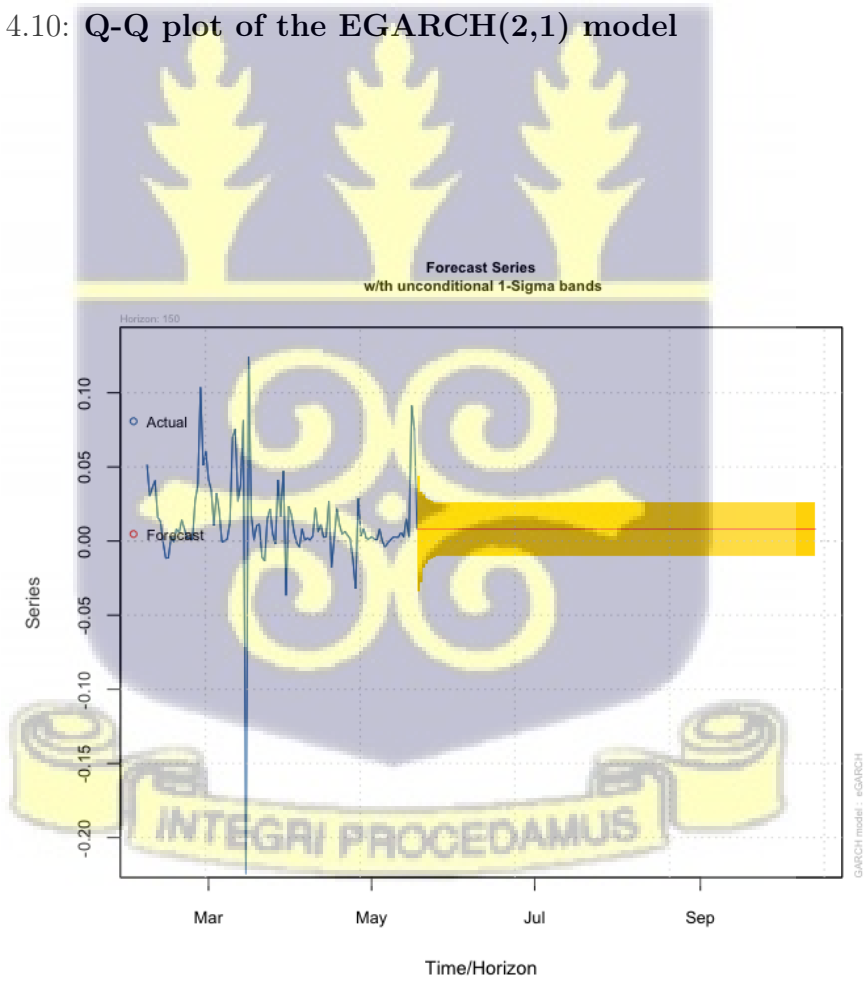


Figure 4.11: **Forecast plot of the EGARCH(2,1) model**

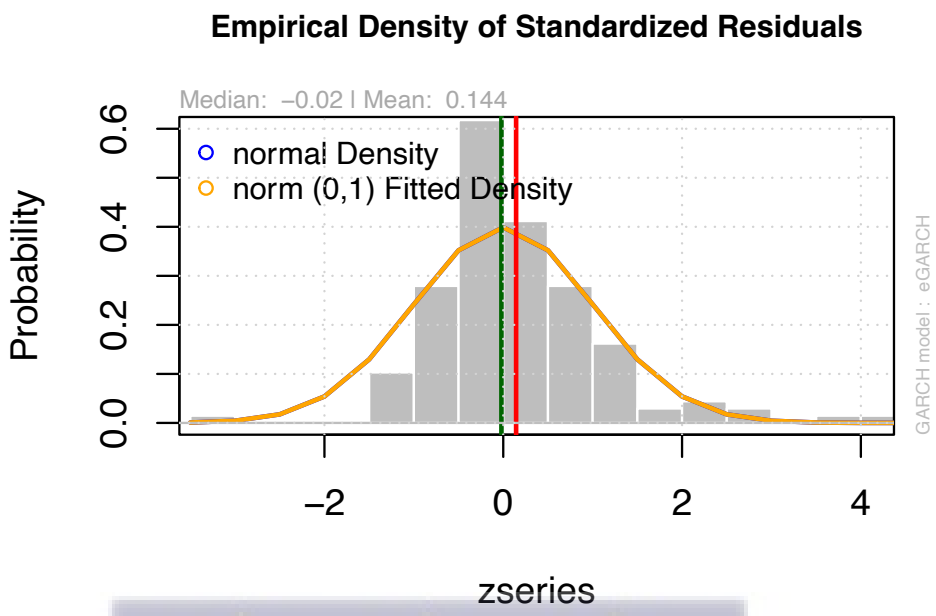


Figure 4.12: **Density plot of the EGARCH(2,1)) model**

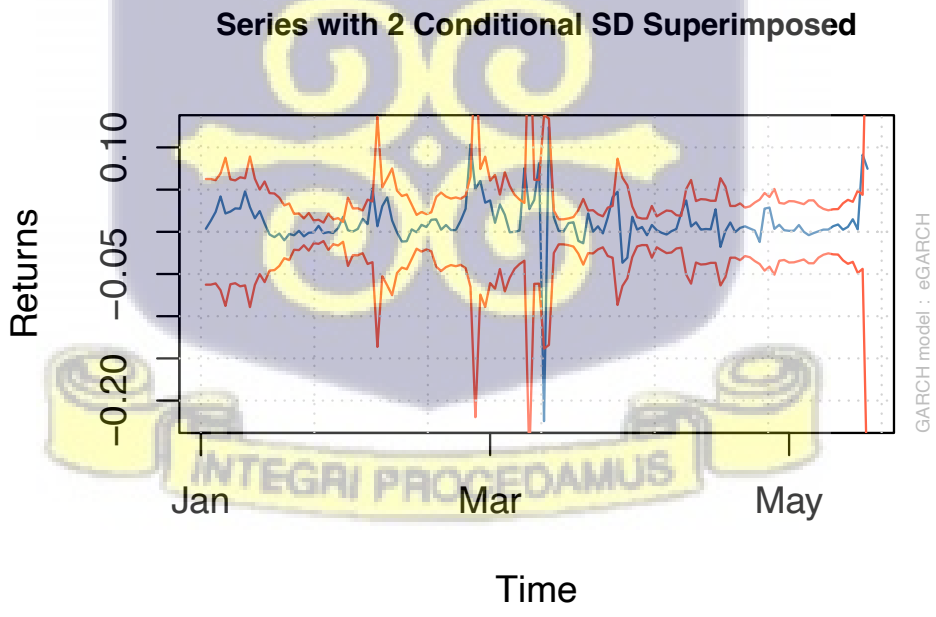


Figure 4.13: **Plot of the volatilities from the EGARCH(2,1) model**

Figures 4.13 displayed the volatility clustering property clearly with regular spikes for the EGARCH model.

In summary, no serial correlation was shown by the residuals since the test for the ARCH effect showed no sign of heteroscedasticity in the residuals. This means that the model was efficient for the task. Checks on the Q-Q plot of the residuals revealed that the model closely seem to be following a normal distribution. After all these checks, it was now clear that the chosen model possessed the desirable properties to fit our data.

4.3.4 Bayesian VAR Model Selection Based on Information Criterion and Lag Selection

The Bayesian Vector Autoregression (BVAR) aspect considered other macroeconomic variables that were somehow related or could determine the exchange rates. The variables considered were GDP and inflation. The stationary plot of the three variables after differencing the data once is shown in Figure 4.14.

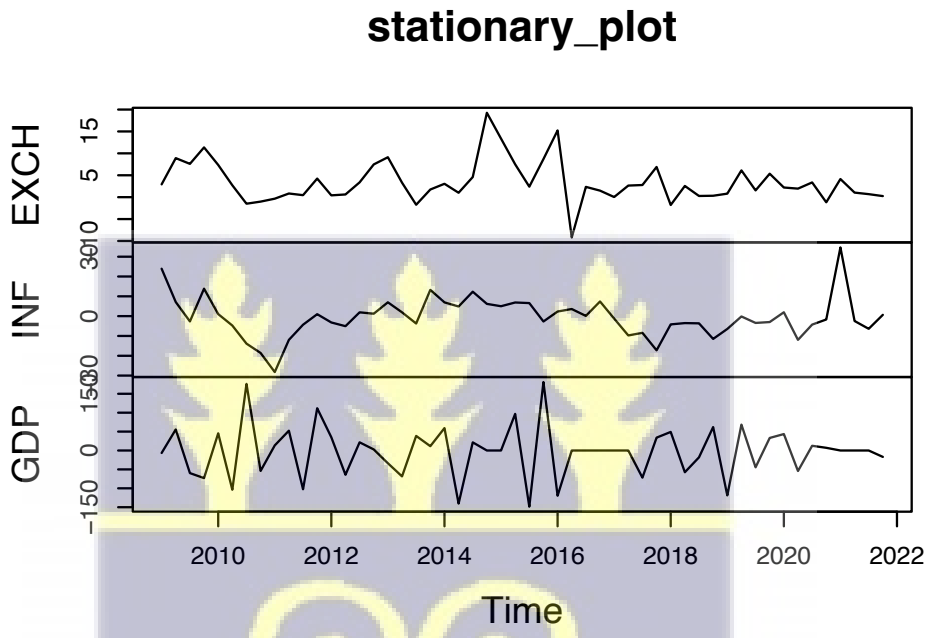


Figure 4.14: **Time series plot of the variables**

The Bayesian VAR had the least values for the AIC, BIC and H-Q when considered at lag 2. That means that the BVAR would give good estimates of our model parameters which could eventually produce good forecast values. The results were presented in Table 4.18.

The Bayesian estimates of the coefficients of the exchange rate, GDP and inflation with their corresponding deviations from their averages were presented in the following equations and matrices. With very minute standard deviation, it was obvious the estimates of the mean values of the parameters were all significant.

Table 4.18: BVAR model selection based on AIC, BIC and H-Q with their corresponding lags

Lag	AIC	BIC	H-Q
1	-9.986	-8.297	-9.375
2	-165.869	-162.491	-164.648
3	-178.543	-173.476	-176.711

The general equation for the Bayesian VAR with two lags is

$$X_i \sim EXCH_{t-1} + INF_{t-1} + GDP_{t-1} + EXCH_{t-2} + INF_{t-2} + GDP_{t-2} + \alpha_i$$

The following derivation can also be made from the general equation

$$EXCH_t = \alpha_1 + \beta_{11}EXCH_{t-1} + \beta_{12}INF_{t-1} + \beta_{13}GDP_{t-1} + \beta_{14}EXCH_{t-2} + \beta_{15}INF_{t-2} + \beta_{16}GDP_{t-2}$$

$$GDP_t = \alpha_2 + \beta_{21}GDP_{t-1} + \beta_{22}INF_{t-1} + \beta_{23}EXCH_{t-1} + \beta_{24}GDP_{t-2} + \beta_{25}INF_{t-2} + \beta_{26}EXCH_{t-2}$$

$$INF_t = \alpha_3 + \beta_{31}INF_{t-1} + \beta_{32}EXCH_{t-1} + \beta_{33}GDP_{t-1} + \beta_{34}INF_{t-2} + \beta_{35}EXCH_{t-2} + \beta_{36}GDP_{t-2}$$

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{pmatrix} \begin{pmatrix} X_{1,\,t-1} \\ X_{2,\,t-1} \\ X_{3,\,t-1} \end{pmatrix} + \begin{pmatrix} \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{24} & \beta_{25} & \beta_{26} \\ \beta_{34} & \beta_{35} & \beta_{36} \end{pmatrix} \begin{pmatrix} X_{1,\,t-2} \\ X_{2,\,t-2} \\ X_{3,\,t-2} \end{pmatrix}$$

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} 0.2021 \\ 0.5882 \\ 0.6495 \end{pmatrix} + \begin{pmatrix} 0.9746 & 0.0239 & 0.0028 \\ 0.3291 & -0.0899 & -0.2846 \\ 1.5001 & 0.4929 & 0.0242 \end{pmatrix} \begin{pmatrix} X_{1,\,t-1} \\ X_{2,\,t-1} \\ X_{3,\,t-1} \end{pmatrix} +$$

$$\begin{pmatrix} 0.0369 & -0.0231 & -0.0094 \\ 0.2811 & 0.1432 & 0.2358 \\ -0.5546 & -0.4258 & 0.0428 \end{pmatrix} \begin{pmatrix} X_{1,\,t-2} \\ X_{2,\,t-2} \\ X_{3,\,t-2} \end{pmatrix}$$

Again, the covariance matrix which shows how the variables (Exchange rate, GDP and Inflation) are related and by how much they vary from each other is presented. Results from the matrix made it known that the GDP varies hugely by itself and when paired with the other variables. The next variable that varies significantly is inflation. The exchange rate variable did not show much variability. So the estimate for the exchange rate component was perceived to be good and may produce reliable forecasts as desired.

A strong correlation is observed between the exchange rate and inflation variables with a value of 0.1416. These two variables are significant in that, a unit increase in the exchange rate between the Cedi/US dollar causes the a unit rise in inflation. This could be because more cedis are required to obtain a dollar for the payment of duties in clearing imported goods into the country. These rates and charges incurred in obtaining these goods or services are then passed onto the final consumers of the finished goods or service.

As goods and services become expensive to acquire in the country, producers of goods and services cannot bear the cost of production of these goods or services hence leading to a reduction in production or a total shut down in production. Therefore, resulting in a drop in GDP in the entire country. This is noted from the covariance matrix with a variance of 0.1450 and from Figure 4.19, that an inverse relation exist between the exchange rate and GDP variables. Whenever the rate of exchange goes up, GDP drops to lower levels.

$$\sum^{-1} = \begin{pmatrix} 0.2367 & 0.1416 & 0.1450 \\ & 8.2920 & -3.1005 \\ & & 24.2380 \end{pmatrix}$$

A histogram of the Bayesian estimates from the posterior draws were displayed in Figure 4.15. That for inflation and exchange rate had a bell shape and looked symmetric, which means data for these two variables are normally distributed. Some of the histogram of the estimates were skewed and which are not significant.

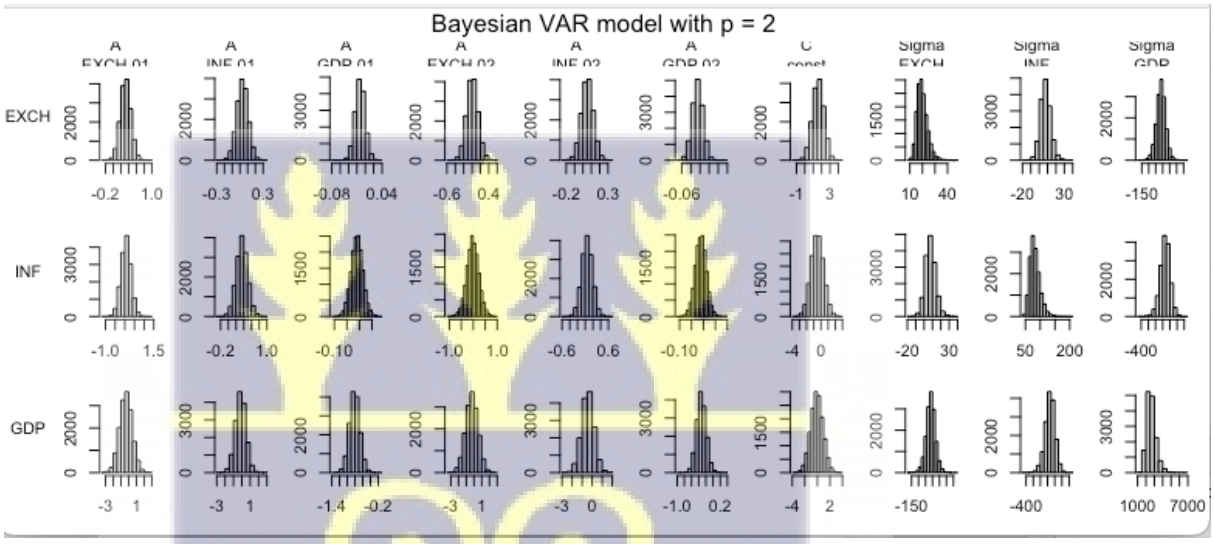


Figure 4.15: **Plot of bvar estimate**

The trace plots of the Bayesian estimates converges and follows the same distribution (normally distributed) across iterations as shown in Figure 4.16.

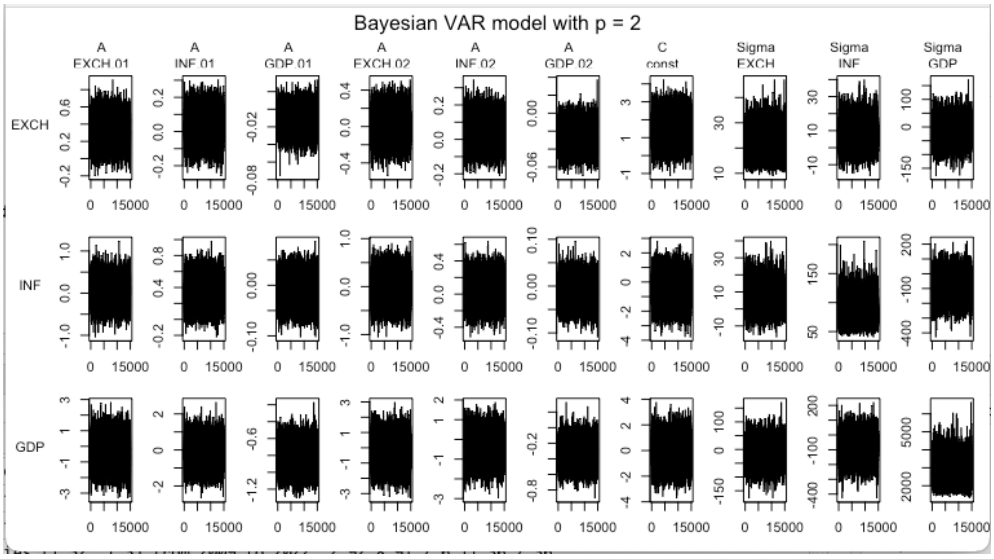


Figure 4.16: **Trace plot of Bvar estimate**

The forecast plot of the three variables revealed that the exchange rate rose slightly and remain stable over the years ahead, while that of the GDP rose sharply then fell and remained stable for the subsequent years. That of inflation did not experience any changes as depicted at the tail ends of the plots in Figure 4.17. It suggests the forecast for both the exchange rate and inflation variables were exact or consistent.

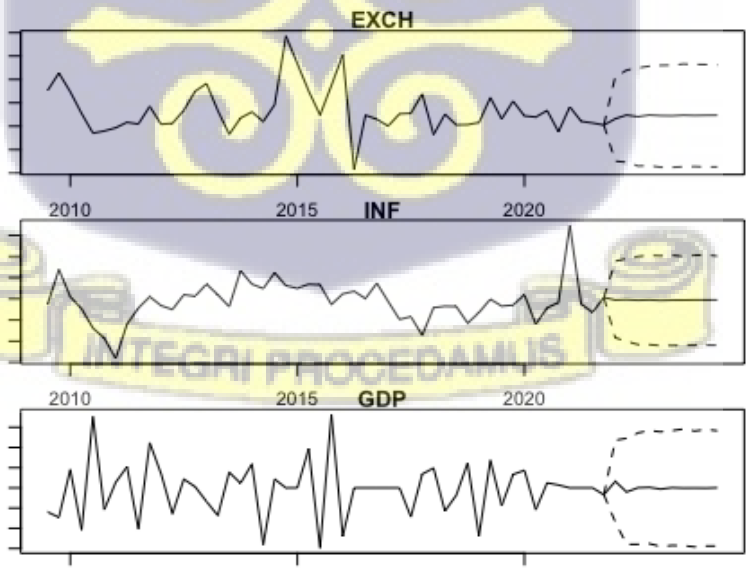


Figure 4.17: **Forecast plot of the variables**

The forecast errors of the Bayesian VAR remained low compared to the previous models discussed. Table 4.19 made it clear that the forecast for the exchange rate were more accurate than that of inflation and GDP.

Table 4.19: Error measures of $BVAR$ forecast at 50% quantile

Variable	RMSE	MAE	MAPE
EXCH	0.017	0.012	0.0734
INF	0.1502	0.1341	0.3872
GDP	0.1502	0.1341	0.3872

A shock from inflation caused the exchange rate to rise sharply roughly by 0.8 and subsequently fell to low levels in the period ahead as shown in Figure 4.18. This tells us that the two variables are positively related.

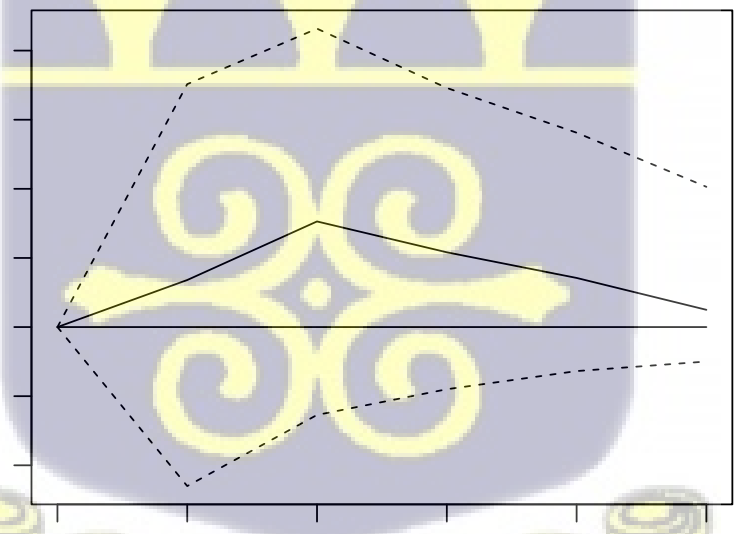


Figure 4.18: **Plot of the impulse response of exchange rate to an inflationary shock**

Meanwhile a shock from GDP caused the exchange rate to fall significantly to lower levels. In other words, GDP negatively impacted the exchange rate from Figure 4.19

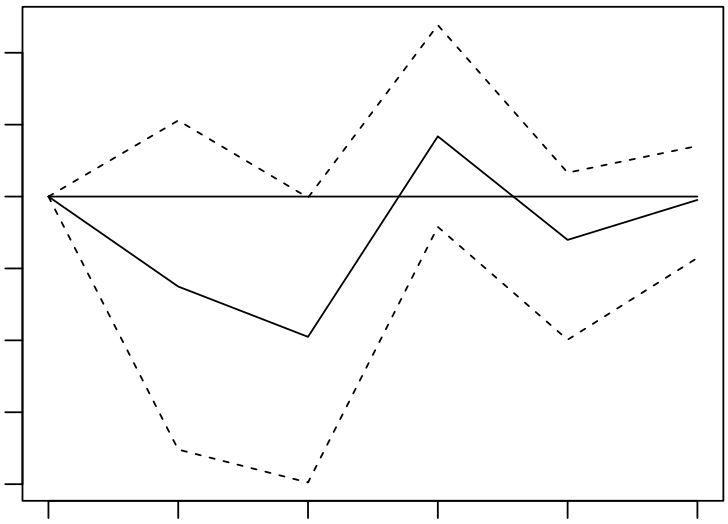


Figure 4.19: **Plot of the impulse response of exchange rate to a GDP shock**

Observing Figure 4.20 which was the error variance analysis of the exchange rate, it could be seen that the exchange rate was relatively influenced by its own shock in the early periods, but in subsequent years the variance of inflation and GDP impacted the exchange rate at minimal levels.

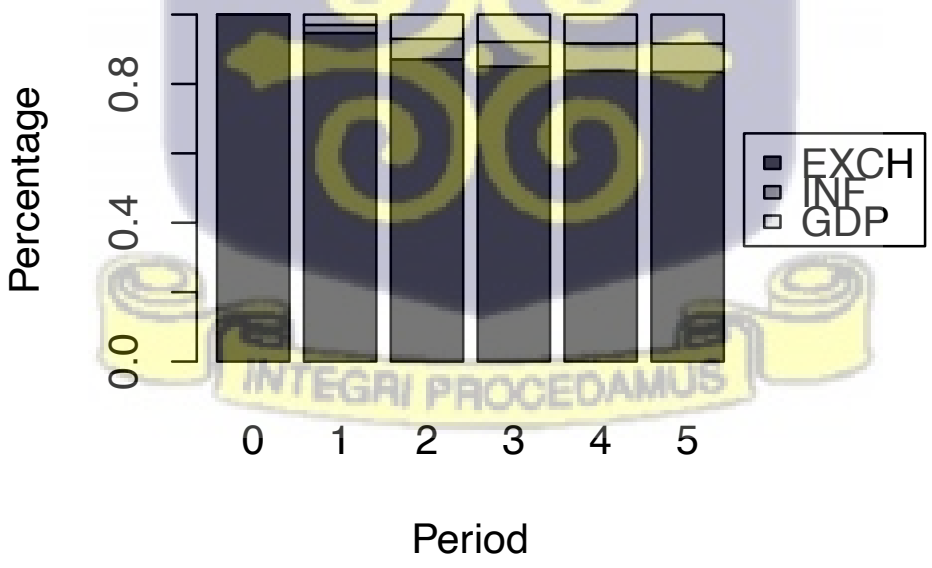


Figure 4.20: **Plot of the forecast error decomposition of the variables**

A final comparison of all the time series models fitted to the data was done using their RMSE and MAE. The Bayesian VAR being a multivariate time series model stood out amongst the models fitted. With an RMSE of 0.017 and MAE of 0.012 respectively, the Bayesian VAR proved to be a good model for forecasting the data.

Table 4.20: Error measures of the forecast values of the three models

Model	RMSE	MAE
BVAR	0.017	0.012
ARIMA	4.557	3.472
GARCH	0.0399	0.0811
EGARCH	0.0326	0.0175

4.3.5 Conclusion

The analysis of the data was done by fitting univariate time series models (ARIMA, Standard GARCH and EGARCH) and a multivariate time series model (Bayesian VAR) since the data was a time series one.

The Standard GARCH and Exponential GARCH models were modelled to the residuals of the ARIMA model. This is so because our data is characterized by volatilities and it was prudent to use the volatility models becuase they had the ability to capture the variances within the residuals. EGARCH (2,1) gave optimal results in capturing the variances in the residuals.

It was evident that the BVAR was capable of fitting the data and was consistent in forecasting the data. Above all, its forecasting errors were much lower compared to the univariate ones.

On this basis, it could be said that multivariate time series model are good in modeling and forecasting time series data than univariate time series models.

Chapter 5

SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This part comprises the summary, conclusion and recommendation from the entire study.

5.2 Summary of the Study

This study examined both univariate time series model (ARIMA, Standard GARCH , EGARCH) and a multivariate model (Bayesian VAR) used in time series analysis. The univariate models were used on the residuals of the ARIMA whilst the multivariate model was used on the actual data.

Data from a secondary source (the Central Bank) was used in the analysis. The data collected covered a 14-year period (January, 2008 to March, 2022). It was segmented into the training and testing sets. The training set (made up of 80%) was used to estimate the parameters of the model while the testing set (made up of 20%) was used in forecasting. An initial plot of the data showed that it was not stationary hence it was made stationary after first differencing and dispalyed using a time series plot. The stationarity of the data was then verified by the KPSS and ADF tests.

In the initial analysis of the data, series of ARIMA models were obtained whereby ARIMA (0,1,1) was chosen to be the optimal model to fit the data.

After obtaining the first difference of the data, serial correlation was still present in the data. A follow up analysis was done by using the ARCH-type models specifically the Standard GARCH and Exponential GARCH models were fitted to the residuals of the ARIMA model.

These ARCH-type models were used because, they possess the abilities of capturing the variances among the residuals by concerntrating on its assymmetries. The Standard GARCH captures only the positive assymmetries within the residuals whilst the EGARCH is able to capture both the positve and negative assymmetries but concentrates more on the negative assymmetries wihtin the data.

In the process, series of the EGARCH models were derived and EGARCH (2,1) was selected as the optimal model with an AIC of -5.0361 and BIC of -4.8648 respectively. Its model parameters were significant with an RMSE and MAE of 0.0326 and 0.0175 respectively.

A multivariate model specifically the Bayesian VAR was also fitted to the actual data to give a broader view on the analysis by assessing the impact of other macroeconomic variables on the exchange rate data and forecasting the variables as well.

The Bayesian VAR tend out to out-perform its univariate counterparts by producing consistent forecast of the data at a 50% quantile and showing the impact of other variables on the exchange rate data. Its forecast errors remained low drastically with an RMSE of O.O17 compared to the univariate models. Obtaining two lags of the data saw its model parameters being significant with a mean of 0.2367 and 0.055.

5.3 Conclusion

The purpose of the study was to obtain the best model that would help capture the uncertain nature of the exchange rate, and use this model in projection. The study specifically seek to:

1. Derive the best forecasting model of the conditional variance.
2. Use a Bayesian VAR to fit and predict the exchange rate data
3. Examine the relationship between the exchange rate and other economic variables.

The conclusions were:

- Based on the analysis of the data, the EGARCH (2,1) was slightly better than the standard GARCH in terms of forecasting the conditional variance with an RMSE of 0.0326.
- The Bayesian VAR was the best model for forecasting the exchange rate as it produced forecast values that were consistent with the actual values. Its forecast error was the least among the models used with an RMSE of 0.017.
- From the analysis of the data, the exchange rate and the inflation variables were positively correlated to each other whilst the exchange rate was inversely correlated to the GDP. GDP and Inflation were also inversely correlated to each other. It means that, a unit increase in the exchange rate causes inflation to rise by a unit vice versa for GDP. This could be because more cedis are required to obtain a dollar for the payment of duties in clearing imported goods into the country. These rates and charges incurred in obtaining these goods or services are then passed onto the final consumers of the finished goods or service.

As goods and services become expensive to acquire in the country, producers of goods and services cannot bear the cost of production of these goods or services hence leading to a reduction in production or a total shut down in production.

5.4 Recommendations

It is recommended that the Bayesian VAR should be used in econometric analysis since it produces better forecast values and its forecast errors are minimal.

Our model EGARCH (2,1) used in fitting the variance of the residuals from the ARIMA model was fairly good but it could be improved with the use of a more complex model such as the GJR-GARCH or quadratic GARCH (QGARCH), which could improve the precision of the estimates.

Also, as the whole idea of the Bayesian inference depends on the choice of prior used, one could consider using a prior (eg. the Gamma prior) that seems to match the data best.

Further research can be done using other varieties of the Bayesian VAR, for instance, the Bayesian Global Vector Autoregression (BGVAR), Structural VAR (SVAR) or the Stochastic Search Variable Selection (SSVS).
Moreover, the two key economic variables (ie. exchange rate and inflation) identified in this study need to be closely monitored in the Ghanaian economy.

Bibliography

Agyemeng, M. (2021). *Impact of macroeconomic variables on exchange rate movements in Ghana*. PhD thesis, University of Ghana.

Aidoo, E. (2010). Modelling and forecasting inflation rates in Ghana: An application of sarima models.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723.

Akumbobe, R. (2015). *Modeling Exchange Rate Volatility Using Univariate GARCH Models-A Case Study Of The Cedi/Dollar Exchange Rate*. PhD thesis.

Alagidede, P., Baah-Boateng, W., and Nketiah-Amponsah, E. (2013). The Ghanaian economy: An overview. *Ghanaian Journal of Economics*, 1(1):4–34.

Anderson, T. W. (1978). Repeated measurements on autoregressive processes. *Journal of the American Statistical Association*, 73(362):371–378.

Anno-Kwakye, R. (2017). *Bayesian Hierarchical Model With Classification And Regression Tree In Predicting Loan Default*. PhD thesis, University of Ghana.

Antwi, S., Issah, M., Patience, A., and Antwi, S. (2020). The effect of macroeconomic variables on exchange rate: Evidence from Ghana. *Cogent Economics & Finance*, 8(1):182–483.

Appleyard, D. R. and Suresh, S. G. (2016). Crown rule, home charges, and uk-india terms of trade. *The Indian Economic Journal*, 63(4):632–653.

Argy, V. (2013). *International macroeconomics: Theory and policy*. Routledge.

Aryeetey, E. and Fenny, A. P. (2017). Economic growth in Ghana. *The Economy of Ghana Sixty Years After Independence*, 45.

Aviah, S. (2018). *Volatility Modelling of Prices of Cocoa Beans Traded on the Intercontinenteal Exchange (ICE*. PhD thesis, University of Ghana.

Banerjee, A., Marcellino, M., and Masten, I. (2005). Leading indicators for euro-area inflation and gdp growth. *Oxford Bulletin of Economics and Statistics*, 67:785–813.

Bartram, S. M., Dufey, G., and Frenkel, M. R. (2005). A primer on the exposure of non-financial corporations to foreign exchange rate risk. *Journal of Multinational Financial Management*, 15(4):394–413.

Bergman, J. J., Noble, J. S., McGarvey, R. G., and Bradley, R. L. (2017). A bayesian approach to demand forecasting for new equipment programs. *Robotics and Computer-Integrated Manufacturing*, 47:17–21.

Bhasin, V. (2004). Dynamic inter-links among the exchange rate, price level and terms of trade in a managed floating exchange rate system: The case of Ghana.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

Caraiani, P. (2010). Forecasting romanian gdp using a bvar model. *Romanian Journal of Economic Forecasting*, 13(4):76–87.

Carriero, A., Mumtaz, H., and Theophilopoulou, A. (2012). Forecasting fiscal variables using bayesian vars with constant and drifting coefficients. *Unpublished (PDF)*.

Chowdhury, M. N. M., Uddin, M. J., and Islam, M. S. (2014). An econometric analysis of the determinants of foreign exchange reserves in Bangladesh. *Journal of World Economic Research*, 3(6):72–82.

Commandeur, J. J., Koopman, S. J., and Ooms, M. (2011). Statistical software for state space methods. *Journal of Statistical Software*, 41:1–18.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366):427–431.

Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, pages 1057–1072.

Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric reviews*, 3(1):1–100.

Dordunoo, C. K. (1994). The structure and policy implications of a macroeconomic model of Ghana. *World Development*, 22(8):1243–1251.

Edwards, W., Lindman, H., and Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological review*, 70(3):193.

Emmanuel, A. (2015). *Modeling GDP Using Vector Autoregressive (VAR) Models: An Empirical Evidence from Ghana*. PhD thesis, University of Ghana.

Engle, R. F. (1990). Stock volatility and the crash of'87: Discussion. *The Review of Financial Studies*, 3(1):103–106.

Gottman, J. M. and Ringland, J. T. (1981). The analysis of dominance and bidirectionality in social development. *Child Development*, pages 393–412.

Hannan, E. J. and Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2):190–195.

Harvey, A. and Phillips, G. (1982). The estimation of regression models with time-varying parameters. In *Games, economic dynamics, and time series analysis*, pages 306–321. Springer.

Iyoboyi, M. and Muftau, O. (2014). Impact of exchange rate depreciation on the balance of payments: Empirical evidence from Nigeria. *Cogent Economics & Finance*, 2(1):923–323.

Jebuni, C. D. and Accra, G. (2006). Challenges and prospects of relaxing export supply constraints in Africa. *centre for policy analysis Accra, Ghana*.

Jebuni, C. D., Sowa, N. K., and Tutu, K. A. (1991). *Exchange rate policy and macroeconomic performance in Ghana*. Initiatives, Nairobi, Kenya.

Jones, R. H. (1980). Maximum likelihood fitting of arma models to time series with missing observations. *Technometrics*, 22(3):389–395.

Kenny, G., Meyler, A., and Quinn, T. (1998). Bayesian var models for forecasting irish inflation.

Khordehfrosh Dilmaghani, A. and Tehranchian, A. M. (2015). The impact of monetary policies on the exchange rate: A GMM approach. *Iranian Economic Review*, 19(2):177–191.

Kim, C.-J. and Nelson, C. R. (1989). The time-varying-parameter model for modeling changing conditional variance: The case of the lucas hypothesis. *Journal of Business & Economic Statistics*, 7(4):433–440.

Kozumi, H. and Polasek, W. (2000). A bayesian semiparametric analysis of arch models. In *Optimization, Dynamics, and Economic Analysis*, pages 389–400. Springer.

Krol, R. (2010). Forecasting state tax revenue: A bayesian vector autoregression approach. *California State University, Department of Economics*.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1):159–178.

Laryea, A. A. and Senadza, B. (2017). Trade and exchange rate policies since independence and prospects for the future. *The Economy of Ghana Sixty Years After Independence*, pages 103–116.

Laryea, E., Ntow-Gyamfi, M., and Alu, A. A. (2016). Nonperforming loans and bank profitability: evidence from an emerging market. *African Journal of Economic and Management Studies*.

Liddle, A. R. (2007). Information criteria for astrophysical model selection. *Monthly Notices of the Royal Astronomical Society: Letters*, 377(1):L74–L78.

Lim, C. M. and Sek, S. K. (2013). Comparing the performances of garch-type models in capturing the stock market volatility in malaysia. *Procedia Economics and Finance*, 5:478–487.

Litterman, R. B. (1986). Forecasting with bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics*, 4(1):25–38.

Lopez, J. A. (2001). Evaluating the predictive accuracy of volatility models. *Journal of forecasting*, 20(2):87–109.

Mankiw, N. G. and Reis, R. (2006). Pervasive stickiness. *American Economic Review*, 96(2):164–169.

Marcellino, M., Stock, J. H., and Watson, M. W. (2003). Macroeconomic forecasting in the euro area: Country specific versus area-wide information. *European Economic Review*, 47(1):1–18.

Mark, N. C. (1988). Time-varying betas and risk premia in the pricing of forward foreign exchange contracts. *Journal of Financial Economics*, 22(2):335–354.

McGrayne, S. B. (2011). *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, & Emerged Triumphant from Two Centuries of C.* Yale University Press.

Mordi, C. N. (2006). Challenges of exchange rate volatility in economic management in Nigeria.

Muchiri, M. (2017). *Effect of inflation and interest rates on foreign exchange rates in Kenya.* PhD thesis, University of Nairobi.

Mumuni, Z. and Owusu-Afriyie, E. (2004). Determinants of the cedi/dollar rate of exchange in Ghana: A monetary approach. *A Bank of Ghana Working Paper.*

Mussa, M. (1986). Nominal exchange rate regimes and the behavior of real exchange rates: Evidence and implications. In *Carnegie-Rochester Conference series on public policy*, volume 25, pages 117–214. Elsevier.

Naceur, S. B., Ghazouani, S., and Omran, M. (2007). The determinants of stock market development in the Middle-Eastern and North African region. *Managerial Finance.*

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the econometric society*, pages 347–370.

Nerlove, M. and Diebold, F. X. (1990). Autoregressive and moving-average time-series processes. In *Time Series and Statistics*, pages 25–35. Springer.

Nortey, E. N., Mbeah-Baiden, B., Dasah, J. B., and Mettle, F. O. (2014). Modelling rates of inflation in Ghana: an application of arch models. *Current Research Journal of Economic Theory*, 6(2):16–21.

Odutayo, A. (2015). Conditional development: Ghana crippled by structural adjustment programmes. *Retrieved February*, 8:2016.

Okoth, M. N. (2013). *The effect of interest rate and inflation rate on exchange rates in Kenya*. PhD thesis, University of Nairobi.

Owusu-Antwi, G., Antwi, J., Ashong, J. D., and Owusu-Peprah, N. T. (2016). Evidence on the co-integration of the determinants of foreign direct investment in Ghana. *Journal of Finance and Economics*, 4(2):23–45.

Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.

Ramos, F. F. (1996). Forecasting market shares using var and bvar models: A comparison of their forecasting performance. *University of Porto, Faculty of Economics, Porto*.

Sammy, O. F. (2018). *Modelling rates of inflation in Kenya : An application of GARCH and EGARCH Models*. PhD thesis.

Schwartz, S. J., Zamboanga, B. L., and Jarvis, L. H. (2007). Ethnic identity and acculturation in hispanic early adolescents: mediated relationships to academic grades, prosocial behaviors, and externalizing symptoms. *Cultural Diversity and Ethnic Minority Psychology*, 13(4):364.

Senadza, B. and Diaba, D. D. (2017). Effect of exchange rate volatility on trade in Sub-Saharan Africa. *Journal of African Trade*, 4(1):20–36.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, pages 1–48.

Sparks, J. J. and Yurova, Y. V. (2006). Comparative performance of arima and arch/garch models on time series of daily equity prices for large companies. *2006 SWDSI proceedings*, pages 563–573.

Spiegelhalter, D. J. (2004). Incorporating bayesian ideas into health-care evaluation.

Spulbar, C. and Birau, R. (2023). Monetary policy transmission mechanism in romania over the period 2001 to 2012: a bvar analysis. In *Research Anthology on Macroeconomics and the Achievement of Global Stability*, pages 203–215. IGI Global.

Stock, J. H. and Watson, M. W. (2017). Twenty years of time series econometrics in ten pictures. *Journal of Economic Perspectives*, 31(2):59–86.

Stockman, A. C. (1987). The equilibrium approach to exchange rates. *FRB Richmond Economic Review*, 73(2):12–30.

Suhartono, S. (2005). A comparative study of forecasting models for trend and seasonal time series does complex model always yield better forecast than simple models. *Jurnal Teknik Industri*, 7(1):22–30.

Takada, T. (2008). Asymptotic and qualitative performance of non-parametric density estimators: a comparative study. *The Econometrics Journal*, 11(3):573–592.

Takaendesa, P. (2006). *The behaviour and fundamental determinants of the real exchange rate in South Africa*. PhD thesis, Rhodes University.

Tsay, R. S. (2005). *Analysis of financial time series*. John wiley & sons.

Tutberidze, D. and Japaridze, D. (2021). A bayesian approach to vector autoregressive model estimation and forecasting with unbalanced data sets. *ECOFORUM*, 10(3).

Tutuani, B. (2015). *Analysis of Spatial Differentials of Income Inequality in Ghana: Application of Bayesian Estimation*. PhD thesis, University of Ghana.

Yao, F. (2011). Forecasting several north dakota microeconomic variables. a bayesian vector auto-regression approach. *West Virginia University, Economics Department, Morgan Town*.

Yule, G. U. (1927). Vii. on a method of investigating periodicities disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 226(636):267–298.

Zellner, A., Palm, F., et al. (1974). Time series analysis and simultaneous equation econometric models. *Journal of econometrics*, 2(1):17–54.

Zivot, E. (2009). Practical issues in the analysis of univariate garch models. In *Handbook of financial time series*, pages 113–155. Springer.

APPENDIX

Appendix A

Table 5.1: Model estimates for exchange rate

Variable	Pairs	Lag	Mean	Std.Dev	2.5%	50%	97.5%
Exchange rate	EXCH	1	0.97461	0.0014398	0.6386	0.9747	1.2986
	INF	1	0.02399	0.000172	-0.0164	0.0238	0.0646
	GDP	1	0.002819	0.000161	-0.0328	-0.0027	0.0392
	EXCH	2	0.036913	0.0014874	-0.02933	0.0354	0.3756
	INF	2	-0.023085	0.000162	-0.06207	-0.0231	0.0163
	GDP	2	-0.009429	0.000157	-0.0448	-0.0094	-0.0259
	Const		0.20214	0.00552	-0.8603	0.2046	1.2592
Inflation	EXCH	1	0.4929	0.00483	-0.6391	0.4911	1.6198
	INF	1	1.50006	0.00092	1.2751	1.5009	1.7247
	GDP	1	-0.02419	0.00074	-0.1530	0.0227	0.2015
	EXCH	2	-0.42580	0.00493	-1.5836	-0.4295	0.7337
	INF	2	-0.55460	0.0009	-0.7847	-0.5549	-0.3271
	GDP	2	0.04282	0.00072	-0.1309	0.0425	0.2162
	Const		0.64948	0.00808	-1.2111	0.6571	2.5269
GDP	EXCH	1	-0.28463	0.0056	-1.5749	-0.0.2916	1.0148
	INF	1	-0.08996	0.00152	-0.4610	-0.0886	0.2753
	GDP	1	0.32908	0.00124	0.0321	0.3298	0.6243
	EXCH	2	-0.23578	0.00576	-1.1009	0.2409	1.5377
	INF	2	0.14315	0.00157	-0.2319	0.1413	0.5244
	GDP	2	0.28107	0.00122	-0.0086	0.2787	0.5771
	Const		0.58820	0.00801	-1.3491	0.5767	2.5028

Forecast tables of the various models

Table 5.2: Variance-Covariance Matrix for exchange rate,inflation and GDP

Variables	Mean	Std.Dev	2.5%	50%	97.5%
EXCH-EXCH	0.2367	0.00053	0.1513	0.2287	0.3683
EXCH-INF	0.1416	0.00235	-0.3168	0.1351	0.6247
EXCH-GDP	0.1450	0.00435	-0.6772	0.1373	0.9982
INF-EXCH	0.1416	0.00235	-0.3168	0.1351	0.6247
INF-INF	8.2920	0.01819	5.3198	8.0035	12.9110
INF-GDP	-3.1005	0.02212	-8.2618	-2.9384	1.1688
GDP-EXCH	0.1450	0.00435	-0.6772	0.1373	0.9982
GDP-INF	-3.1005	0.022115	-8.2618	-2.9384	1.1688
GDP-GDP	24.2380	0.0526	15.4840	23.4171	37.8168

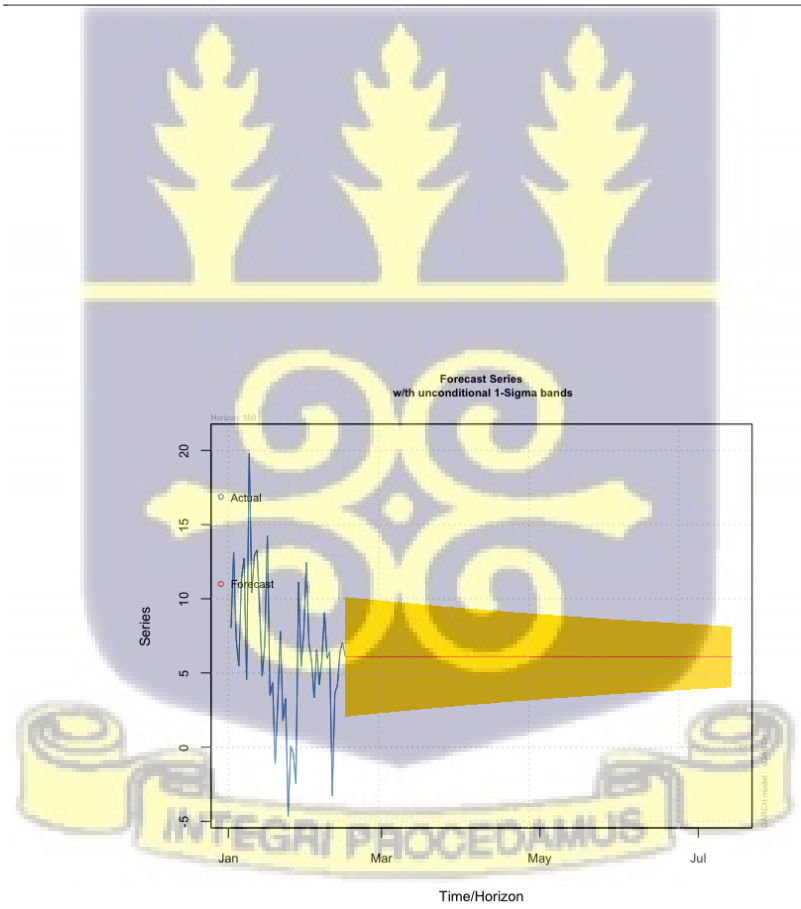


Figure 5.1: **Forecast plot of the standard GARCH model**

Table 5.3: Bayesian forecast for exchange rate at 95% Credible Interval

Year	Quarter	Forecast	Min	Max
2019	Q1	15.2092	13.2090	16.2809
2019	Q2	15.5482	12.4229	17.7250
2019	Q3	15.8545	11.5505	19.2556
2019	Q4	16.3963	10.5662	20.8979
2020	Q1	16.2044	9.4657	22.6566
2020	Q2	16.8887	8.2487	24.5316
2020	Q3	17.0657	6.9159	26.5225
2020	Q4	17.1841	5.4674	28.6291
2021	Q1	17.2266	3.9032	30.8514
2021	Q2	17.2421	2.2226	33.1901
2021	Q3	17.5191	0.4243	35.6464
2021	Q4	17.8242	-1.4934	38.2222

Table 5.4: Bayesian quantile forecast for exchange rate

Year	Quarter	Actual	Year	Quarter	2.5%forecast	50%forecast	97.5%forecast
2019	Q1	15.2092	2019	Q1	13.6791	14.7350	15.7643
2019	Q2	15.5482	2019	Q2	13.4878	15.1006	16.6865
2019	Q3	15.8545	2019	Q3	13.2624	15.4649	17.6421
2019	Q4	16.3963	2019	Q4	13.0681	15.8609	18.5918
2020	Q1	16.2044	2020	Q1	12.8777	16.2489	19.6765
2020	Q2	16.8887	2020	Q2	12.6453	16.6453	20.7741
2020	Q3	17.0657	2020	Q3	12.4386	17.0299	21.9178
2020	Q4	17.1841	2020	Q4	12.2324	17.4304	23.0619
2021	Q1	17.2266	2021	Q1	11.9679	17.8595	24.3967
2021	Q2	17.2421	2021	Q2	11.7124	18.2673	25.6781
2021	Q3	17.5191	2021	Q3	11.4136	18.6666	27.1415
2021	Q4	17.8242	2021	Q4	11.1701	19.0875	28.6785

Table 5.5: EGARCH (2,1) forecast of the conditional variance for exchange rate at 95% Confidence Interval

Year	Quarter	Forecast	Min	Max
2019	Q1	0.0361	0.0361	0.0361
2019	Q2	0.0262	0.0176	0.3318
2019	Q3	0.0216	0.058	0.4131
2019	Q4	0.0193	0.0025	0.2732
2020	Q1	0.0181	0.0032	0.4408
2020	Q2	0.0175	0.0025	0.3205
2020	Q3	0.0172	0.0035	0.3672
2020	Q4	0.0170	0.0034	0.4712
2021	Q1	0.0169	0.0035	0.3220
2021	Q2	0.0186	0.0054	0.4373

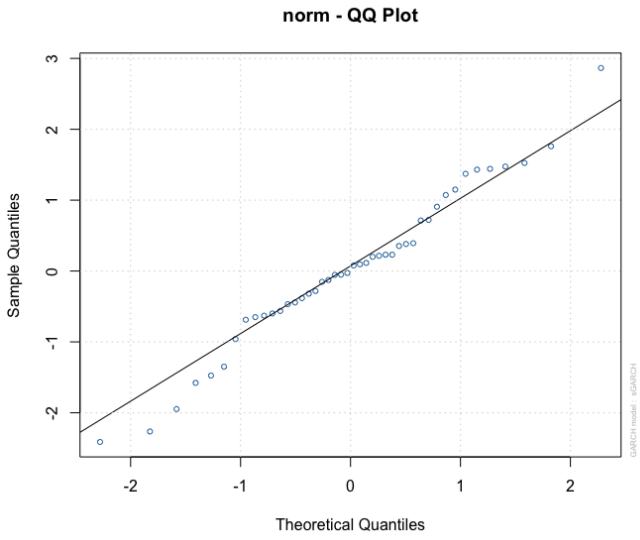


Figure 5.2: **Quantile plot of the standard GARCH model**

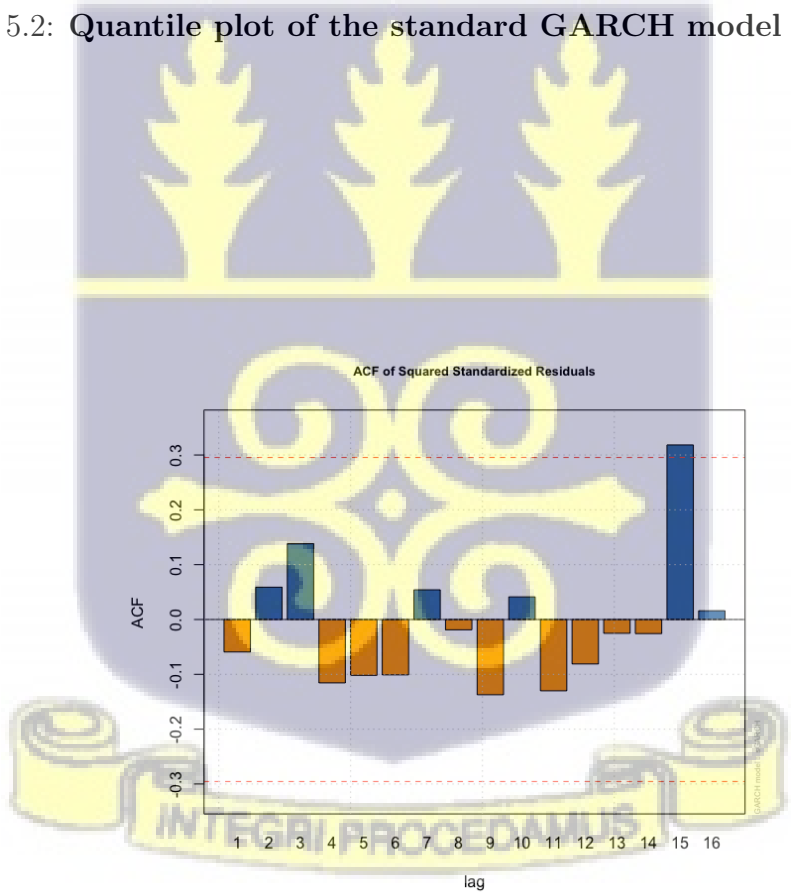


Figure 5.3: **ACF plot of the squared standardized residuals of the standard GARCH**

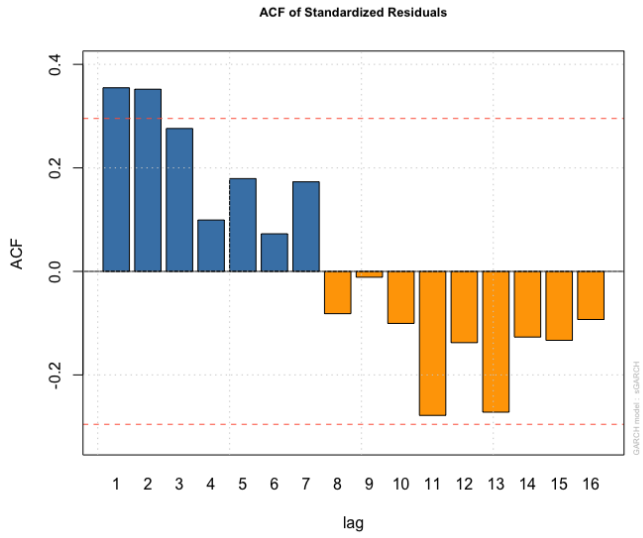


Figure 5.4: **ACF plot of the standardized residuals from the standard GARCH model**

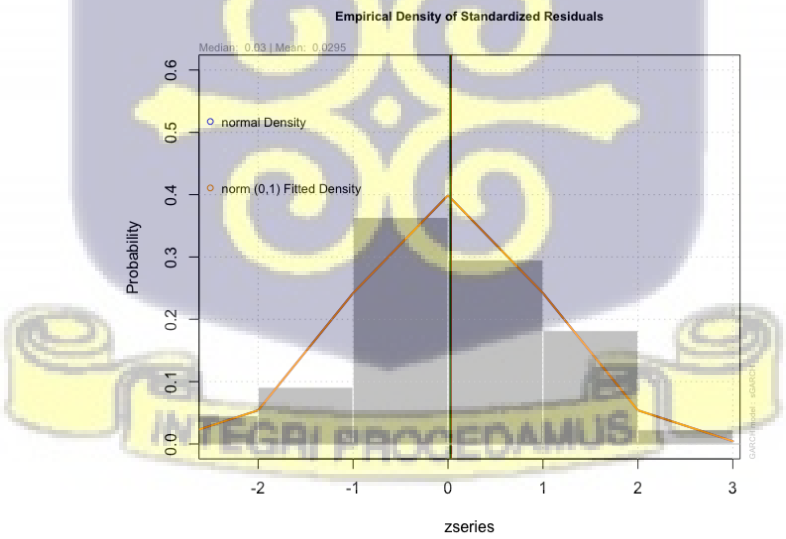


Figure 5.5: **Density plot of the standard GARCH**

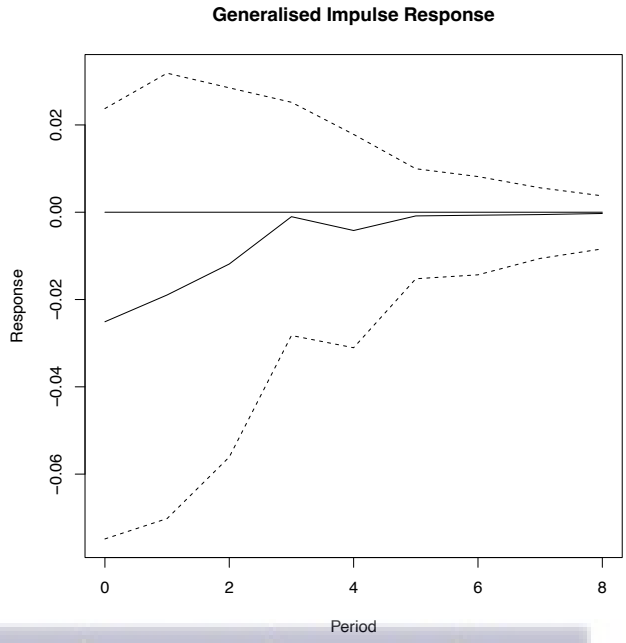


Figure 5.6: **Plot of the Genralized Impulse Response of GDP on Inflation**

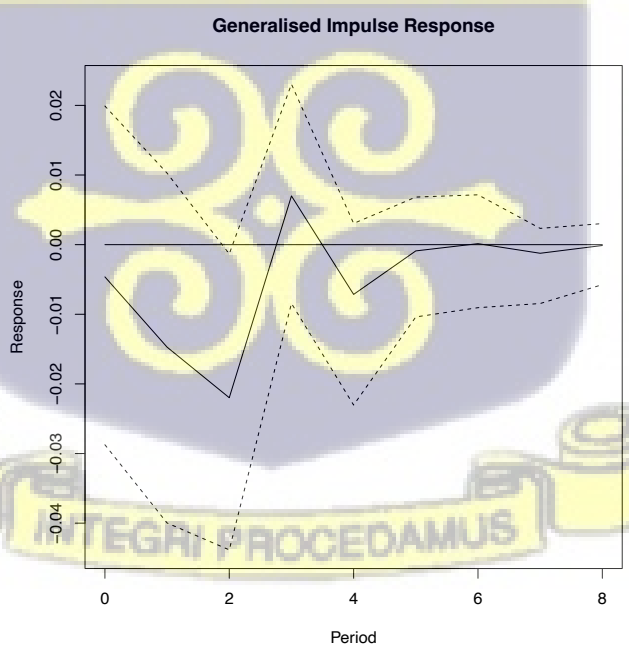


Figure 5.7: **Plot of the Generalized Impulse Response of GDP on Exchange rate**

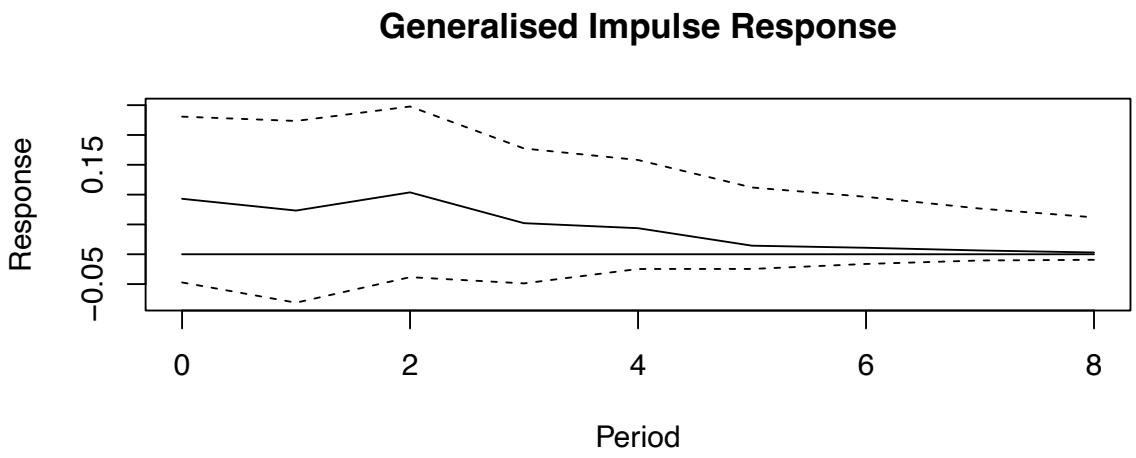


Figure 5.8: **Plot of the Generalized Impulse Response of Inflation on Exchange rate**

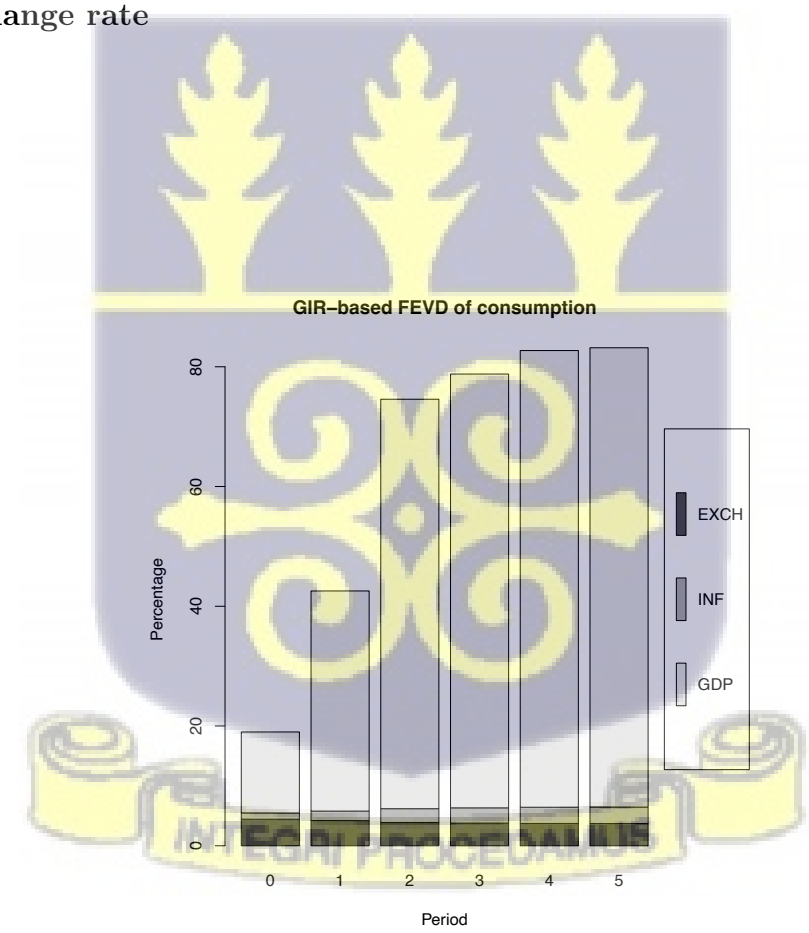


Figure 5.9: **Plot of the impact of Generalized Impulse Response and Forecast Error Variance Decomposition on exchange rate**

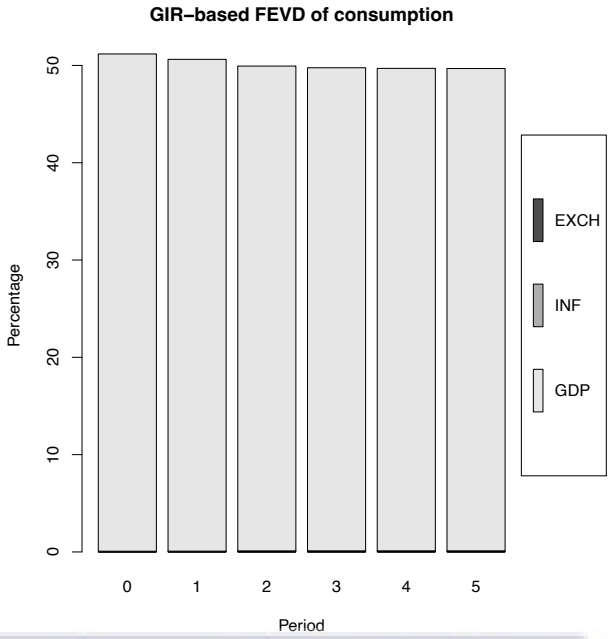


Figure 5.10: **Plot of the impact of Generalized Impulse Response and Forecast Error Variance Decomposition on GDP**

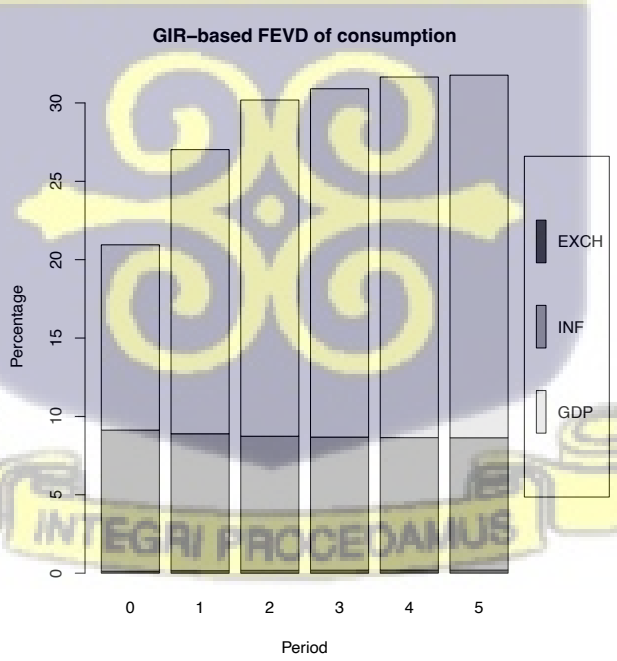


Figure 5.11: **Plot of the impact of Generalized Impulse Response and Forecast Error Variance Decomposition on Inflation**