UNIVERSITY OF NAIROBI
SCHOOL OF MATHEMATICS

MODELLING EXCHANGE RATE VOLATILITY OF KES/USD USING GARCH FAMILY MODELS

SYLVIA ATIENO OMOLO

I574/84354/2012

SUBMITTED IN PARTIAL FULFILLMENT OF THE
MASTER OF SCIENCE DEGREE IN ACTUARIAL SCIENCE

July 14, 2014

DECLARATION

This project is my own work and has not been presented in any other University.

Signature...

SYLVIA ATIENO OMOLO

This project has been submitted with my approval as a University Supervisor.

Signature..

CAROLE OGUTU
DEPARTMENT OF MATHEMATICS

ACKNOWLEDGEMENT

First, I would like to thank God for the far He has brought me, the journey has not been easy.

My sincerest gratitude to my Supervisor Carole Ogutu for her guidance and availability throughout this process. I am indebted for your dedication.

Many thanks to Dr.Mwaniki for his very important input and suggestions, Professor J.A.M Otieno for his invaluable input at the inception of this project and Professor Patrick Weke, Dr. Nelson Owuor and Idah Orowe for their insight and advice.

Lastly, I would like to sincerely thank my Family for constantly cheering me on and being my best fans and critics. Without your support, encouragement and prayers, it would not have been possible.

LIST OF ABBREVIATIONS

ACF   Auto Correlation Function

AIC - Akaike Information Criterion

ADF - Augmented Dickey-Fuller

AGARCH -   Asymmetric Generalized Auto Regressive Conditional Heteroscedasticity

AR - Auto Regressive

ARCH - Auto Regressive Conditional Heteroscedasticity

ARIMA - Auto Regressive Integrated Moving Average

ARMA -   Auto Regressive Moving Average

APARCH -   Asymmetric Power Auto Regressive Conditional Heteroscedasticity

BDT - Bangladeshi Taka

CBK -   Central Bank of Kenya

CGARCH -   Component Generalized Auto Regressive Conditional Heteroscedasticity

EGARCH -   Exponential Generalized Auto Regressive Conditional Heteroscedasticity

FOREX -   Foreign Exchange

GARCH -   Generalized Auto Regressive Conditional Heteroscedasticity

GED -   Generalized Error Distribution

GNP -   Gross National Product

GJR GARCH -   Glosten-Jagannathan-Runkle Generalized Auto Regressive Conditional Heteroscedasticity

JB -   Jarque Bera

JPY -   Japanese Yen

KES -   Kenya Shillings

LERMS -   Liberalized Exchange Rate Management System

LM -   Lagrange Multiplier

LPR -   Log Price Relative

MA -   Moving Average

MAE -   Mean Absolute Error

MAPE -   Mean Absolute Percentage Error

PAC -   Partial Auto Correlation

PACF - Partial Auto Correlation Function

PARCH - Power Auto Regressive Conditional Heteroscedasticity

PCA -   Principal Components Analysis

PP -   Phillip Perron

QMLE -   Quasi Maximum Likelihood Estimation

RMSE -   Root Mean Square Error

TARCH -   Threshold Auto Regressive Conditional Heteroscedasticity

TGARCH -   Threshold Generalized Regressive Conditional Heteroscedasticity

UK -   United Kingdom

USD -   United States Dollar

Contents

ABSTRACT

Investors, Policy makers, Governments etc. are all consumers of exchange rates data and thus exchange rate volatility is of great interest to them. In this paper, the importance of exchange rate volatility modelling has been brought out in the problem statement and the significance of the study.

The paper thus seeks to study the KES/USD exchange rate volatility using GARCH family models. Symmetric and asymmetric models have been used to capture volatility characteristics of exchange rates on data spanning 23/10/1993 to 21/03/2014 and a comparison is made under four different conditional distribution assumptions i. e. Normal, Skewed Normal, Student-t and Skewed Student-t distributions.

Investigations conducted showed that the asymmetric EGARCH model with skewed student-t distribution emerges as the best based on AIC and log-likelihood. It can account for the asymmetry in the exchange rate return series while the skewed student-t accounts for some of the skewness and leptokurtosis.

The data is further split into four sub-periods and volatility characteristics analyzed per period. Each period shows a significant occurrence of high fluctuations in the exchange rate.

Chapter 1

INTRODUCTION

1.1 Exchange rates

The exchange rate between two currencies represents the value of one country's currency in terms of another country's currency. Different countries use different mechanisms to keep their currency stable by identifying an exchange rate regime that best suits their economy. Essentially, there are two categories i.e. systems with fixed exchange rates and those with flexible exchange rates. In between the two extremes lies a range of intermediate systems.

In the fixed regime (hard peg), the exchange rate is usually a political decision where a government attempts to keep the value of its currency constant against another, known as the anchor currency. For the flexible exchange rate regime, prices are usually determined by market forces in accordance with demand and supply (Abdalla, 2011). They can either be independent whereby no attempt is made to maintain a particular rate and exchange rate is determined by the market. They can also be managed floating, whereby the exchange rates are market determined but the country's monetary authority can intervene when the need arises so as to prevent undue fluctuations.

Floating exchange rates are expected to be more volatile as they are free to fluctuate. The volatility in exchange rates results in increase of exchange rate risk and adversely affects the international trade and investment decisions(Kamal et al. ,2011).

1.2 Evolution of Kenya's exchange rate system

Between 1969 and 2009, the KES/USD exchange rate transitioned from fixed to crawling to floating eras. Between 1966 and 1992, Kenya operated a fixed exchange regime and the country had to frequently devalue its currency to reduce the negative effects that real exchange rate volatility had on its economy (Munyoki et al.,2012)

The floating exchange rate system was adopted in 1993; however, there is no available evidence that success has been achieved in realizing the objective for which the foreign exchange market was liberalized (Munyoki et al.,2012)

1.3 Exchange rate volatility

Volatility refers to the spread of all unlikely outcomes of an uncertain variable (Abdalla, 2011). Exchange rate volatility refers to fluctuations in exchange rates over time. Exchange rate fluctuations have received

much attention because it has an influence on inflation, international trade, investment analysis, profitability and risk management among others (Abdalla, 2011). For most financial time series, Bollerslev (1986)'s GARCH (1, 1) model has been found to be sufficient for modelling exchange rate volatility.

Volatility provides an idea of how much the exchange rate will change over a given period and is usually calculated from the standard deviation. Two measures of volatility are the Historical and implied volatility. Historical volatility is the realized volatility over a given time period and can be obtained whereby, with a series of past exchange rates, the standard deviation of the daily price changes can be calculated and these can then be extended to annual volatility. Implied volatility is a forward looking measure calculated from the market participant estimates of what is likely to happen in the future. A financial time series , say $\{Y_t\}$, is often a serially uncorrelated sequence with zero mean , even as it exhibits volatility clustering, suggesting that the conditional variance of Y_t given past returns is not constant.

Many financial crises such as those of Latin America, Southeast Asia and Russian economies stemmed from sudden and unexpected oscillation of foreign exchange thus highlighting the importance of measurement of exchange rate volatility, its forecasting and behavior (Kamal et. al, 2011). Exchange rate movements impact on volume and value of foreign trade and investment in that exchange rate volatility tends to affect imports and exports which further influences a country's balance of payments. Lothian & Taylor(1997) put forth a theory that suggested that import substitution takes place when real exchange rate depreciation happens and relative cost of domestic to foreign goods is lowered. Increased volatility could thus lead to a reduction in direct investment.

Statistically, volatility is often measured as a sample standard deviation i.e.
$$\hat{\sigma} = \sqrt{\frac{1}{N-1} \sum_{t=1}^{N} (r_t - \mu)^2}$$

where
r_t = return on day t
μ = average return over the N days period

NB: At times, the variance σ^2 is also used as a measure of volatility. In this paper, variance is used as the volatility.

Researchers have documented empirical regularities regarding financial time series. Some of the stylized facts about exchange rate volatility include:

- Leverage Effects
  This gives the relationship between returns and volatility. Black(1976) noted that for stock returns, volatility was higher after negative shocks than after positive shocks of the same magnitude. He alluded this asymmetry to leverage effects. In financial markets, depreciation of currency is usually followed by higher volatility i.e. negative returns seem to be more important predictors of volatility than positive returns Engle (1993). This phenomenon is known as leverage effect.

- Mean reversion
  This states that there is a level of volatility to which volatility will eventually return, it does not persist forever. After an increase or a decrease, volatility returns to this mean level no matter how long it takes.

- Volatility clustering
  Large and small values in the log-returns tend to occur in clusters i.e. the large changes tend to

be followed by large changes and small changes tend to be followed by small changes (Mandelbrot, 1963). When volatility is high it is likely to remain high and when it is low it is likely to remain low. This feature reflects on the fact that news is clustered over time (Engle, 2004).

- Persistence
  If today's return has a large effect on the forecast variance many periods in the future, volatility is said to be persistent. Persistence makes it take a while for volatility to recover to the normal level.

- Fat tails
  Comparing normal distribution with financial time series such as exchange rate returns, fatter tails are observed, also known as excess kurtosis. Kurtosis is the standardized fourth moment of a normal distribution and usually has a value of 3 but for financial time series, a value above 3 is usually observed. Mandelbrot (1963)

- Non-Normality
  Exchange rates tend to deviate from normality. This can be observed from the histogram plot of the data hence the need to model it using other conditional distributions that can better capture it's skewness and kurtosis.

Usually, information tends to influence the value of one currency in terms of another e.g announcements made by the central bank of a country of their intention to intervene or not, in the markets for their currencies. Further, the macroeconomic performance of an economy's currency also influences exchange rate volatility because no trader wants to hold a given currency for an economy that appears too risky.

Many economic time series exhibit periods of unusually large volatility followed by periods of relative tranquility. In such cases the assumption of constant variance is inappropriate.

Some of the well known models used to estimate exchange rate volatility are the ARCH model by Engle (1982) and the GARCH model by Bollerslev (1986). Baillie & Bollerslev(1989) noted that the GARCH class of models have the ability to capture the volatility dynamics of exchange rates at daily, weekly and monthly frequencies. A lot of work has been done by academic and policy researchers in analyzing sources of exchange rate volatility yet modelling it still remains a challenge. This is partly evidenced by numerous theoretical models of exchange rate determination and several modelling approaches Chipili (2006). The ARCH process by Engle (1982) allows the conditional variance to change over time as a function of past errors with the unconditional variance constant, on the other hand, conventional time series and economic models assume constant variance.

Findings show that ARCH models could efficiently represent typical empirical findings e. g. conditional heteroscedasticity in financial time series. With the collapse of the Bretton Woods system and application of flexible exchange rates in the seventies, ARCH models have become more popular.

GARCH models are useful in capturing the leptokurtic nature of financial time series data as well as volatility clustering and help in modelling the changing conditional variances in time series (Bollerslev, 1986).Further, Brooks & Burke are of the view that lag order (1, 1) is sufficient to capture all the volatility clustering in the data. Least squares model has been used by econometricians to determine how much one variable will change in response to another variable. The basic version of the least squares model assumes that the expected value of all error terms when squared is the same at any given point i.e. homoscedasticity. More econometricians are now being asked to forecast and analyze the size of the errors

of the model. ARCH/GARCH models can do this. Data that has variances of the error terms not being equal i. e. larger for some points than others suffer from heteroscedasticity. Assumptions underlying the GARCH model are that the time series under consideration must exhibit serial correlation as well as long term dependence.

1.4   Problem statement

The Kenya shilling has registered mixed performances against the USD. The fluctuations ranged between 35 in 1994 when the Kenya shilling was strongest and 105 in 2011 when it was at its weakest. This shows how volatile the KES/USD exchange rate has been. These fluctuations tend to increase exchange rate risk. This risk is what needs to be examined and perhaps quantified.

In the Monetary Policy for Fiscal year 2011/2012, CBK attributed the appreciation of the Kenya shilling to the tight monetary policy stance adopted by the Monetary Policy Committee. Also, the disbursements from the IMF through the Enhanced Credit Facility programme towards the end of 2011 and the disbursement of USD 600 million syndicated loan to the government provided a further cushion to the Kenya shilling.

Exchange rate volatility affects policy makers as well as investors hence the need to study volatility which can aid in financial decision making. These kinds of fluctuations present real dangers to economic stability. A weak shilling makes Kenyan goods and services cheaper in the international market but makes imports more expensive so exporters benefit while importers lose. Conversely, a strong shilling makes Kenyan goods and services expensive in the international market and makes imports more affordable. Volatility thus brings about questioning of the stability of financial markets. Depending on the direction a government wants its economy to make, they may create monetary policies that either appreciate or depreciate their currency. Policy makers essentially rely on volatility estimations so as to enable them make decisions on what direction the currency should take.

In the Kenyan Market, Maana et al.(2010) and Latifa et al. (2013) concluded that the GARCH(1,1) model was sufficient for estimation of volatility in various currency exchange rates, the KES/USD being one of them. GARCH(1,1) is normally applied with the normal distribution. This model however, can only capture some of the skewness and fat tails in financial time series. Also financial time series tend to exhibit non-normality hence the need to explore other conditional distributional forms for the error term that can better capture the characteristics of the data.

The study thus attempts to model the volatility behavior of KES/USD exchange rates which is crucial to diverse groups of people such as governments, importers, exporters, corporate decision makers, investors, e.t.c. by obtaining a model that can better capture all the characteristics of the KES/USD exchange rates data.

1.5 Significance of the study

The study seeks to model the variance of KES/USD exchange rates. It will facilitate investigation of exchange rate risk which can be an indicator of vulnerability in the economy. This will enable the government manage the volatility, at least in the short run as well as provide reliable models for policy makers to help them anticipate possible vulnerabilities of financial markets and the economy and to analyze and forecast volatility of exchange rates that can guide the central bank to intervene in the market when the need arises. Further, it will bring about an understanding of the behavior of exchange rates as well as an attempt to explain the sources of these movements as well as fluctuations.

It would also inform various consumers of exchange rate data and guide policy makers on currency risk and how to handle a currency crisis if and/when it occurs. Importers, exporters, currency traders among others will be able to assess market conditions before making their transactions.

The Kenyan market has many foreign currency activities thus the study could be of great benefit to investors. They can use this information to make decisions on future investments from the observed patterns of exchange rate volatility considering that investors' confidence to invest in a particular country is inversely related to high volatilities in exchange rates. Policy makers on the other hand can use information on short term volatility to enable them decide on intervention policies if need be. The Government will also be able to manage the exchange rate volatility atleast in the short run. Kenya as a developing country would benefit greatly from this kind of research which hasn't been explored much in spite of the fact that exchange rate fluctuations have a huge contribution to inflation.

1.6 Objectives

1.6.1 Overall objective

To model the exchange rate volatility of the KES/USD using GARCH family models.

1.6.2 Specific objectives

This research will therefore seek to answer the following underlying objectives:

1. To identify the volatility characteristic of the KES/USD exchange rate.

2. To build a volatility model that captures the volatility of KES/USD exchange rate.

The rest of this paper is organized as follows: chapter 2 is the literature review, chapter 3 highlights the methodology, chapter 4 is data analysis and results and finally chapter 5 gives the conclusions and recommendations.

Chapter 2

LITERATURE REVIEW

Bollerslev (1986) acknowledges the usefulness of the ARCH process in modelling several different economic phenomena. However, he noted in most of those applications , the introduction of a rather arbitrary linear declining lag structure in the conditional variance equation to take account of the long memory typically found in empirical work because estimation of a totally free lag distribution often would lead to violation of the non-negativity constraints. He then came up with GARCH, a more general class of processes which allowed for a more flexible lag structure which also permitted a more parsimonious description.

In empirical applications of the ARCH model, a relatively long lag in the conditional variance equation is often required and to avoid problems with negative variance parameter estimates a fixed lag structure is typically imposed. This led to an interest in the extension of the ARCH class of models to allow for both a longer memory and a more flexible lag structure. The GARCH (p, q) process is then described followed by the simplest in this class, the GARCH (1, 1).

The usefulness of the autocorrelation and partial autocorrelation functions in identifying and checking time series behavior in the conditional variance equation of the GARCH is shown. He also considers maximum likelihood estimation of the GARCH regression which differed from Engle (1982) because of the inclusion of the recursive part. The maximum likelihood estimate is found to be strongly consistent and asymptotically normal. He also found that a general test for GARCH (p, q) is not feasible as in Godfrey (1978). Bollerslev (1986) then used the GARCH process to model the uncertainty of inflation, an unobservable economic variable of major importance. He proceeded to explain the rate of growth in the implicit GNP deflator in the US in terms of its own past. The model was estimated on quarterly data, a total of 143 observations, using ordinary least squares. The model was found to be stationary and none of the first ten autocorrelations or partial autocorrelations for ϵ_t were significant at 5% level. Findings showed that the GARCH (1, 1) model provided a slightly better fit than the ARCH (8) model in Engle and Kraft (1983)as well as exhibiting a more reasonable lag structure.

Engle(1982) noted that for conventional econometric models, the conditional variance did not depend upon its own past. He thus proposed the ARCH model which was able to capture the idea that today's variance does depend on its past as well as the non-constant nature of the one period forecast variance. He then went ahead to define the likelihood function for the ARCH processes as well as a formulation for the general ARCH process. He also found that the parameter α had to satisfy the non-negativity constraints as well as some stationary conditions. Finally, he used the ARCH model to estimate the means and variances of inflation in the UK.

Latifa et al. (2013) model heteroscedasticity in foreign exchange for US, UK, Euro and Japanese Yen data using GARCH models. Monthly averages for the various currency exchange rates were collected for the period from January 2001 to December 2010, a total of 120 observations per foreign currency. The period was chosen because of the two major milestones that the country underwent i. e. Election period followed by the post election violence in 2007/2008. Their major aim was to study how these events affected the performance of the afore mentioned currencies.

Their findings show uncertainty of exchange rates between 2001 to 2005 , gaining relative stability upto 2008 where the shilling became weaker against the foreign currencies due to post election violence. They noted that the Kenyan economy undergoes a cycle of about five years approximately, from one election period to the next. The GARCH (1,1) model was found to be adequate, thus confirming the work of Bollerslev (1992). They recommended further research on whether Non-linear (N-GARCH) can be appropriate for modelling this kind of data.

Rotich(2014) models USDKES,EURKES and GBPKES exchange rate volatility using the EGARCH model under the assumption of both normal and student-t distribution for comparison purposes. He notes that the student-t EGARCH is more favourable compared to the normal distribution because of evidence of the heavy tailed nature of financial time series. According to him, the normal GARCH model could neither explain the entire fat tail nature of the data nor could it explain the asymmetric responses. He then goes ahead to describe the EGARCH model as well as to give a specification of the two error distributions i. e. normal and student-t.

From his results, the series showed volatility characteristics of returns, non-normality of the return series and presence of ARCH effects. There was also evidence of leverage effects whereby good news produced more volatility than bad news. The GBPKES and the EURKES were both fitted by an EGARCH(1,1) model while USDKES was fitted by an AR(1)/EGARCH(1,1) with $\epsilon_t \approx t(0, 1, \upsilon)$

Awogbemi and Alagbe (2011) examine the volatility of the Naira/USD and Naira/UK pound sterling exchange rates in Nigeria using the GARCH model. They use monthly exchange rates spanning the period 2006-2010. According to them, the main objective of an exchange rate policy is to determine an appropriate exchange rate and ensure its stability and over the years, efforts put by the Nigerian Government to achieve this have not yielded positive results. They thus sought to build a forecasting model that would adequately capture the volatility of Nigerian exchange rate return series using GARCH model and the outcome of their research was to assist the government to manage the exposure of the exchange rate volatility in the short run, inform investors on future behavior of exchange rates thus helping them in decision making and help end users of volatility models such as importers, exporters, etc.

The rate of returns was employed to study the currency exchange rate. From the ADF and PP tests, the exchange rates were found to contain a unit root (non-stationarity). For the normality test, the Jacque Bera test showed that the exchange rate returns series departed from normality. From their summary statistics of exchange rate returns, Naira/USD is leptokurtic and negatively skewed with relative to normal distribution while Naira/UK pound sterling is platykurtic and positively skewed. They established the presence of persistent volatility in return series and suggested that the persistence could be due to the import dependence, inadequate supply of foreign exchange by the Central bank of Nigeria and activities of foreign dealers and parallel market. They suggested that further studies should be done using higher frequency data and other different volatility models and that the impact of government intervention should also be investigated.

Baba Insah (2013) in his paper, investigates the presence and nature of real exchange rate volatility in the Ghanaian economy propelled by the fact that over the years there have been intermittent spikes in volatility which indicated that Ghana's international competitiveness had deteriorated over the period of the study. The exchange rates considered in the study is the log of the real effective exchange rate. The Breusch-Pagan test was used to test for the presence of ARCH (1) effects. As much as ARCH effects were detected using the Breusch-Pagan test, the ARCH model was found not to fit the estimation. However, GARCH (1, 1) model was found to perform well in the modelling of exchange rates. Switch of regime from fixed to floating caused a spike in volatility from 1983 to 1986 but was more minimal between 2001 and 2010.

Abdalla (2011) used daily observations from 19 Arab countries and considered the GARCH approach in modelling exchange rates. Observations in the period 1st January 2000 to 19th November 2011, a total of 4341, were used. The LM test was used to test for heteroscedasticity. GARCH model was then used to investigate the volatility clustering and persistence. The model had only three parameters that allowed for an infinite number of squared errors to influence the current conditional variance (volatility). EGARCH (1, 1) was used to capture leverage effects as GARCH models are poor in capturing these effects.

From the descriptive statistics, skewness and excess kurtosis observed for the daily returns of all currencies indicate a departure from normality. The Jacque-Bera statistics also confirmed this non normality. All the returns displayed positive skewness for 14 currencies and negative skewness for 5 currencies. A highly leptokurtic distribution was also observed for all the series. The ADF test results rejected the null hypothesis of a unit root for all the series. The ARCH LM test was used to test for heteroscedasticity and this led to rejecting the null hypothesis of no ARCH effects in the residual series in the mean equation. GARCH (1, 1) model results suggested that the conditional variance is an explosive process for 10 country currencies and that volatility shocks are persistent for 7 countries. EGARCH (1, 1) results showed leverage effects for all the currencies except for the Jordanian Dinar where negative shocks implied a higher next period conditional variance than positive shocks of the same magnitude. They concluded that exchange rates volatility can be adequately modeled by the class of GARCH models.

Chipili (2006) states that, in spite of the importance of exchange rate volatility in macroeconomics, a study of its sources in Zambia is largely unexplored. This was despite evidence that fluctuations in the kwacha exchange rate have strong effects on inflation". The author considers the few studies known to him on the same, not robust. He sought to examine a large sample of currencies that includes 20 kwacha real and nominal bilateral as well as effective exchange rates chosen on the basis of Zambia's trade structure. The study is conducted over a longer sample period, 1964-2006 using relatively higher frequency data at monthly intervals. He employs both symmetric and asymmetric GARCH models. A PCA was also conducted on the estimated conditional variance to generate a GARCH series that captures a common pattern in all exchange rates that can be used in subsequent empirical work (The author is not aware of any study that has generated PCA exchange rate volatility series from GARCH models and applied it in empirical work as an alternative measure of exchange rate uncertainty.)

GARCH (1, 1), TARCH (1, 1) and EGARCH (1, 1) model specifications are used as they sufficiently capture exchange rate behavior. EGARCH model was used because it allows for oscillatory behavior of the variance and requires no parameter restriction of non-negativity of coefficients like GARCH as it is modeled in log-linear form. The model also captures leverage effect. Non-normality of the exchange rate returns was confirmed by skewness, kurtosis and J-B statistics. Volatility clustering was also visible. The ADF and PP unit root tests confirm non-stationarity of the data.

GARCH model results were robust based on diagnostic tests on residuals that show the absence of serial correlation and no remaining ARCH effects. The TGARCH model on the other hand revealed an asymmetry term that was insignificant at all conventional significance levels suggesting no detectable asymmetry in all exchange rate series except Kwacha/Swiss franc. This evidence of asymmetry suggested that the symmetry imposed by the GARCH (1, 1) model was restrictive. From the results, conditional velocity differed across kwacha exchange rates examined. It can be inferred that,"exchange rates are characterized by different conditional volatility such that imposing a uniform GARCH model specification on all exchange rates may be inappropriate". PCA deals with the variance structure of a set of observed variables through a linear combination of the variables (components). It was conducted on 39 conditional variances generated from GARCH and EGARCH models, split into real and nominal measures. The correlation coefficients indicated very close relationships among conditional volatility in exchange rates with the pattern observed with the Zimbabwe dollar being the only exception. There was evidence in support of EGARCH model as the best fitting model and emphasis was made on the importance of testing the appropriateness of the model specifications as opposed to imposing a uniform GARCH model.

Alam & Rahman (2012) explore the application of GRACH type models to model the BDT/USD exchange rate using daily foreign exchange rate series. The study was based on secondary data with the study period split from July 3rd 2006 to May 13th 2010 for the in-sample data set and May 14th 2010 to April 30th 2012 for the out-of-sample data set. This made a total of 1513 trading days. The Jacque-Berra statistics showed non-normality of BDT/USD exchange rate. The exchange rate series was then converted to the return series because foreign exchange rates are usually non-stationary and quite random hence unsuitable for the study purpose. The ADF and PP tests were then used to check for stationarity of BDT/USD exchange rates return. The exchange rates return series was found to be stationary because the value of the ADF test statistic was less than its critical value at 1 percent level of significance. The PP test gave a similar result. A plot of the BDT/USD exchange rates return series disclosed a slight positive skewness and a higher positive kurtosis. The GARCH, EGARCH, PARCH and TARCH models were then benchmarked with an AR and ARMA model in the study.

Statistical performance measures were calculated to identify the best model in the case of in-sample as well as the out of sample case. The lower the values of RMSE, MAE, MAPE and theil-u, the better the forecasting accuracy of the given model. The trading performance measures such as annualized return, annualized volatility, information ratio and maximum drawdown are used to select the best model. Their findings reveal that ARMA and AR are selected as the best model as per in-sample trading performance outcomes whereas TARCH model is nominated as the best model according to out-of-sample trading performance outcomes without transaction costs. (GARCH model in the case of transaction costs).

Maana et al. (2010) applied the GARCH process in the estimation of volatility of the foreign exchange market in Kenya using daily exchange rates data from January 1993 to December 2006. Currencies used were USD, sterling pound, Euro and Japanese Yen. Data used was obtained from the CBK database and to estimate volatility in exchange rates, logarithm rates returns were used. From the descriptive statistics for exchange rate returns, skewness coefficients were greater than zero indicating that the exchange returns distributions are not normal. The positive skewness coefficients indicate that the distribution of the returns is slightly right skewed implying that depreciation in the exchange rate occur slightly more than appreciation.

Kurtosis coefficients for all currencies returns were greater than three indicating that the underlying distributions of returns are leptokurtic. The Jaque-Berra normality test indicates that the distribution of exchange rate returns for all the currencies have tails which are significantly heavier than that of the

normal distribution. Results for the volatility estimation show that the estimated GARCH (1, 1) models are significant at 5 percent significance level and fit the data well. The plots of the GARCH (1, 1) models revealed decreasing volatility in the exchange rate returns implying relative stability in the exchange rate.

Samsudheen & Shanmugasundaram (2012) attempted to understand the behavior of Indian foreign exchange rate and its volatility characteristics. They used daily exchange rates of rupee/USD spanning a period of forty years: 1st April 1973 to 31st March 2012. They equally wanted to measure the impact of structural changes in the Indian exchange rate system from the pegged exchange rate to the liberalized exchange rate Managed system (LERMS) in 1992 and market determined exchange rate regime in 1993. They used different ARCH family models such as ARCH (1, 1), GARCH (1, 1), EGARCH (1, 1), TARCH (1, 1) etc. for their study. The move by India to the market determined exchange rate is considered as a major structural change in the Indian foreign exchange market and this would lead to frequent volatility in the Indian exchange rate market hence an attempt to better understand the behavior of exchange rates.

To investigate foreign exchange rate volatility they use time series data of daily exchange rates. To measure the impact of structural changes in the exchange rate system, they divided the entire sample period into two sub-periods: Pre-implementation (1973 to 1993) and post implementation (1993 to 2012) denoted as pre and post LERMS. They use percentage daily exchange rates return. The descriptive statistics are analyzed for daily foreign exchange rate series and daily return series of three periods, full sample, Pre LERMS and Post LERMS. From the values obtained for skewness and kurtosis, the foreign exchange rate series data is not normal. This is further confirmed by the Jacque-Berra test. The same results are obtained when daily exchange rate return series data is used. From plots of the exchange rate vs. exchange rate returns series, it was observed that high volatility fluctuation prevailed in post LERMS period than in the pre LERMS period.

To test for stationarity, the ADF and PP tests were used. Results showed that the return series are stationary. Before modelling volatility, they tested for heteroscedasticity. The ARCH LM test used indicated the presence of ARCH effects. The volatility models were then divided into symmetric and asymmetric volatility models. The difference between the two is that in symmetric models the effect of the future volatility is dependent on both magnitude and sign of the underlying return series. The symmetric GARCH (1, 1) model results show that the volatility of the Indian foreign exchange rate is highly persistent in all three periods and is higher in the post LERMS period than in the pre LERMS period. The asymmetric models EGARCH and TGARCH evidenced existence of leverage effect in all the three sample periods more so in the post LERMS period.

Kamal et al. (2011) examine performance of GARCH family models to forecast the volatility behavior of the Pakistani FOREX market. Data used is daily FOREX rates between January 2001 and December 2009. To achieve stationarity, the time series FOREX data was transformed into daily returns. The ADF test was used to check for stationarity. Further results showed that monthly returns of representative exchange rates were independent of serial correlation. Overall results indicated that the EGARCH model remains the best in explaining the volatility behavior of the data. The TARCH model depicted the presence of leverage effect.

Dukich et al (2010) assess the adequacy of GARCH models by considering three exchange rate sequences JPY, Euro, British pound and evaluating how well the GARCH model replicates the empirical nature of

these sequences. Their period of observation was January 4, 1999 to January 4, 2010 with over 2700 trading days. They assessed the assumptions fitting the GARCH model before fitting the model for each of the LPR sequences. Their tests showed that the GARCH model was suitable for fitting this data. Three GARCH models were then fitted: GARCH (1, 1), GARCH (1, 2), GARCH (2, 1). Much as assumptions underlying the GARCH models were satisfied for each of the LPR sequences, none of the GARCH models considered was able to capture the empirical nature of the LPR sequences appropriately. They suggested use of other varieties of GARCH family models to better capture the properties of the LPR series.

Baba Insah (2013) investigated the presence and nature of real exchange rate volatility in the Ghanaian economy. His findings show that GARCH (1, 1) was the right model for modelling exchange rate volatility in Ghana. He started by testing for the presence of ARCH (1) effects using a Breusch-Pagan test. As much as there was evidence of ARCH effects, the ARCH (1) model was not suitable. GARCH (1, 1) on the other hand performed well in modelling the exchange rate volatility.

Zhou(2009) looks at the stylized facts as well as the time varying volatility of S & P 500 stock index. A comparison is made between the ARMA(0,2)/APARCH (1,1) model with the normal, student-t and skew-t distrbution. Asset returns exhibit fat tails as well as asymmetry. Zhou(2009) uses a skewed-t distribution which is able to capture leptokurtosis of asset returns and then models the asymmetry in the conditional variance equation as a non linear GARCH model i.e. the ARMA(0,2)/APARCH (1,1) model with non-gaussian errors. He gives a brief description of the normal, student-t and skewed-t distribution density functions as well as their loglikelihood function which together with AIC and BIC criterions are used to determine the best fitting distribution. The JB test rejected the null hypothesis of normal distribution. To test for stationarity, the ADF and PP tests are used. The transformed return series is found to be stationary. The ACF of the observations and the squared observations show some relevant autocorrelation so an ARMA(p,q) model is used to fit the asset returns. His conclusion is that the ARMA(0,2)-APARCH (1,1) with skew-t distribution is the best fitted model to use for the conditional heteroscedasticity and this is then used to forecast.

Adedayo et al. (2013) use the student-t and GED distribution to model the innovations of the Naira-USD exchange rate. They looked at studies carried out on the Naira exchange rate series such as those by (Olowe, 2009) and(Ezike and Amah, 2011). Some of the shortcomings they saw are that the authors considered monthly data in their investigations and based on this, the series' characteristics were not well captured. Also, the studies only considered one or two exchange rates out of many. Other than that, the studies assumed a normal distribution and did not look at different distributional forms. Lastly, they felt that due to the volatility and asymmetry in exchange rate series, daily data should have been applied to examine these properties. They thus applied various conditional distributions to both symmetric and asymmetric GARCH type models. The conditional distributions considered were student-t and GED. Four exchange rates were selected: Naira-Euro, Naira-British pound, Naira-Japanese Yen and Naira-USD. Data used span between 10/12/2001 and 14/12/2011. An AR(1) model was estimated as the mean equation because in the log return series, autocorrelation was only significant at first lag. They selected an asymmetric GARCH model with t distribution in most cases but for Naira-USD, a GARCH-GED model was specified.

Chaiwat et al. use daily closing price data from Thailand, Malaysia and Singapore exchange rates to study which GARCH(p,q) and one of six different error distributions best fit the data, based on AIC criterion. Because financial time series have been known to exhibit many non-normal characteristics that cannot be captured by the standard GARCH model with gaussian distribution, they saw the need to explore the student t and GED distributions which exhibit heavy tails, as well as the skewed student t

and skewed GED distributions which allow various types of skewness and heavy tails. Their data set was divided into two subsets i.e. in-sample data set used to build up a model for underlying data and the out-of-sample data set used to investigate the performance of volatility forecasting. Their findings show that a GARCH(p,q) model with non-gaussian error distribution tends to provide better out-of-sample forecast performance than a GARCH(p,q) model with normal error distribution. From their simulations, they also find that the MSE given by the best fitted model is not significantly different from that given by the best forecast performance model. The skewed error distributions also outperform other error distributions in terms of out-of-sample volatility forecasting.

Ebru et al.(2013) evaluate volatility forecast performance of asymmetric GARCH models for MIST country currencies vs USD. The MIST countries are: Mexico, Indonesia, South Korea and Turkey which are a new set of emerging market countries . They use monthly exchange rates spanning 1993.01 to 2012.12. They pick the best volatility model as per their study and use this for out-of-sample forecasting. They use various GARCH models (GARCH, EGARCH,APARCH,CGARCH,GJR GARCH, AGARCH) with different distributions because financial time series generally exhibit fat tails and asymmetry and GARCH models with normal distributions usually fail to capture these. Forecast performance for the models is evaluated using MAE, MAPE and Theil-Inequality criteria. Their findings show that for all the models applied, the most suitable distributions were student-t and GED. The best volatility forecasting model was found not to be the same as the best forecasting model. All the three countries exhibited a different volatility model from the forecasting model. They concluded from their estimations that asymmetric GARCH models gave better forecasts than symmetric GARCH models for two of the examined countries.

Chapter 3

METHODOLOGY

3.1 The purely random process

This is a sequence of uncorrelated, identically distributed random variables with zero mean and constant variance and is the simplest type of model used as a building block in many other models. This process is stationary and is also known as white noise, the error process or innovation process. The term white noise arises from the fact that a frequency analysis of the model shows that in analogy with white light all frequencies enter equally. It is denoted by Z_t with zero mean and variance. An alternative notation is ϵ_t. For a purely random process, it can be seen that ϵ_t is strictly stationary as shown below:

$Pr(\epsilon_{t_1} \leq x_1, \epsilon_{t_2} \leq x_2, \ldots, \epsilon_{t_n} \leq x_n)$
$= Pr(\epsilon_{t_1} \leq x_1)Pr(\epsilon_{t_2} \leq x_2)\ldots Pr(\epsilon_{t_n} \leq x_n)$(by independence)
$= Pr(\epsilon_{t_1-k} \leq x_1)\ Pr(\epsilon_{t_2-k} \leq x_2)\ldots Pr(\epsilon_{t_n-k} \leq x_n)$(Identical distributions)
$= Pr(\epsilon_{t_1-k} \leq x_1, \epsilon_{t_2-k} \leq x_2, \ldots, \epsilon_{t_n-k} \leq x_n)$(by independence)

3.2 The random walk

The random walk model is given by

$$X_t = X_{t-1} + \epsilon_t \tag{3.2.1}$$

where ϵ_t denotes a purely random process. Equation (3.2.1) does not form a stationary process, however the first difference, $X_t - X_{t-1}$ forms a stationary series.

Equation (3.2.1) can also be decomposed as:

$X_1 = \epsilon_1$
$X_2 = \epsilon_1 + \epsilon_2$
$\vdots$
$X_t = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_t$

The mean function is given by:
$\mu_t = E(X_t) = E(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_t) = E(\epsilon_1) + E(\epsilon_2) + \cdots + E(\epsilon_t)$

$= 0 + 0 + \cdots + 0$

such that $\mu_t = 0$ for all t

$$Var(X_t) = Var(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_t) = Var(\epsilon_1) + Var(\epsilon_2) + \cdots + Var(\epsilon_t)$$
$$= \sigma_\epsilon^2 + \sigma_\epsilon^2 + \cdots + \sigma_\epsilon^2$$
$$Var(X_t) = t\sigma_\epsilon^2$$

Thus the process variance increases linearly with time.

Investigation of the covariance function:
Suppose $1 \leq t \leq s$, then

$$\gamma_{t,s} = cov(X_t, X_s) = cov(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_t, \epsilon_1 + \epsilon_2 + \cdots + \epsilon_t + \epsilon_{t+1} + \cdots + \epsilon_s)$$

This can be written as

$$\gamma_{t,s} = \sum_{i=1}^{s} \sum_{j=1}^{t} cov(\epsilon_i, \epsilon_j)$$

The covariates are zero unless i=j, in which case they equal $var(\epsilon_i) = \sigma_\epsilon^2$. There are exactly t of these so that $\gamma_{t,s} = t\sigma_\epsilon^2$

3.3   AR processes

A process X_t is said to be an autoregressive process of order p, i.e. AR(p) if it is a weighted sum of the past p values plus a random shock so that

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \cdots + \phi_p X_{t-p} + Z_t \tag{3.3.1}$$

The value at time t depends linearly on the last p values and the model looks like a regression model hence the term autoregression. Using the backward shift operator B such that $BX_t = X_{t-1}$, the $AR(p)$ model may be re-written as $\phi(B)X_t = Z_t$ i.e.
$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \cdots + \phi_p X_{t-p} + Z_t$$

$$Z_t = X_t - \phi_1 B X_t - \phi_2 B^2 X_t - \phi_3 B^3 X_t - \cdots - \phi_p B^p X_t$$

$$Z_t = X_t(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \cdots - \phi_p B^p$$
Hence $Z_t = \phi B X_t$

Where $\phi B = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \cdots - \phi_p B^p$

The first order AR process, AR(1) is

$$X_t = \phi_1 X_{t-1} + Z_t \tag{3.3.2}$$

If $\phi = 1$ then model in equation (3.3.2) reduces to a random walk as in equation (3.2.1) when the model is non stationary. With $|\phi| > 1$ then the series becomes explosive, hence non stationary. However, if $|\phi| < 1$, it can be shown that the process is stationary with acf given by
$\rho_k = \phi^k$ for $k = 0, 1, 2, \ldots$ thus the acf decreases exponentially.

3.4 MA processes

A process is said to be a moving average process of order q, MA (q) if it is a weighted sum of the last random shocks i. e.

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q} \tag{3.4.1}$$

Using the backward shift operator B, it may be written as $X_t = \theta(B)Z_t$ i. e.

$$X_t = Z_t + \theta_1 B Z_t + \theta_2 B^2 Z_t + \cdots + \theta_q B^q Z_t$$
$$X_t = Z_t(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)$$

Hence $X_t = \theta(B)Z_t$ where $\theta B = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ A finite order MA process is stationary for all parameter values. However it is customary to impose a condition on the parameter values of an MA model, known as the invertibility condition to ensure that there is a unique MA model for a given acf. Suppose Z_t and Z'_t are independent purely random processes and that $\theta \in (-1, 1)$, then two MA processes defined by
$X_t = Z_t + \theta_1 Z_{t-1}$ and $X_t = Z'_t + \theta^{-1} Z'_{t-1}$ have exactly the same acf. Thus the polynomial $\theta(B)$ is not uniquely determined by the acf. As a consequence, given a sample acf, it is not possible to estimate a unique MA process from a given set of data without putting some constraint on what is allowed. To resolve this ambiguity, it is usually required that the polynomial $\theta(x)$ has all its roots outside the unit circle. (3.4.1) can be re-written as

$$X_t - \sum_{j \leq 1} \Pi_j X_{t-j} = Z_t$$

For some constraint π_j such that

$$\sum_{j \leq 1} \mid \pi_j \mid < \infty$$

i.e. we can invert the function taking the Z_t sequence to the X_t sequence and recover Z_t from present and past values of X_t by a convergent sum. A more parsimonious model is the ARMA process.
NB:
AR, MA and ARMA processes are used on stationary time series.

3.5 ARMA processes

This is a mixed Autoregressive Moving Average model with p AR terms and q MA terms i. e. ARMA (p, q). This is denoted as $\phi(B)X_t = \theta(B)Z_t$

$$X_t(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) = Z_t(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q \tag{3.5.1}$$

Where $\phi(B)$, $\theta(B)$ are polynomials in B of finite order p, q respectively. (3.5.1) has a unique causal stationary solution provided that the roots of $\theta(x)$ lie outside the unit circle. Using ARMA processes, many real data sets may be approximated in a more parsimonious way by a mixed ARMA model rather than a pure AR or pure MA process.

Many real data sets may be approximated in a more parsimonious way by a mixed ARMA model rather than a pure AR or pure MA process. ARMA models are used to model the conditional expectation of a process given the past, but in an ARMA model, the conditional variance given the past is constant. For, say, daily rates which are constantly changing, at one point they may seem unusually volatile leading one to infer that tomorrows return will also be more variable than usual. An ARMA model cannot capture this because its conditional variance is constant. A better model is thus necessary to model the non constant

volatility. This leads us to the GARCH models which can be used because they allow for conditional changes in the variance.

NB: AR, MA and ARMA processes are used on stationary time series.

3.6 ARIMA processes

In practice, many time series are not stationary hence the AR, MA and ARMA processes cannot be used. To make the series stationary, differencing is used. The first difference may then be differenced a second time to give a second difference and the differencing continues until stationarity is achieved. The first difference is given by

$$X_t - X_{t-1} = X_t - BX_t$$
$$= X_t(1 - B)$$

Second difference:
$$X_t(1 - B) - X_{t-1}(1 - B) = X_t(1 - B) - BX_t(1 - B)$$
$$= X_t(1 - B)(1 - B)$$
$$= X_t(1 - B)^2$$

Third difference:
$$X_t(1 - B)^2 - X_{t-1}(1 - B)^2 = X_t(1 - B)^2 - BX_t(1 - B)^2$$
$$= X_t(1 - B)^2(1 - B)$$
$$= X_t(1 - B)^3$$

Hence the dth difference is given by:
$$X_t(1 - B)^d$$

If the original data series is differenced d times before fitting an ARMA(p, q) process, then the model for the original undifferenced series is said to be an ARIMA (p, d, q) process. The 'I'stands for Integrated and denotes the number of differences taken.

$\phi(B)X_t = \theta(B)Z_t$ is now generalized as $\phi(B)(1 - B)^d X_t = \theta(B)Z_t$

$\phi(B)(1 - B)^d X_t$ becomes the combined AR operator. If we replace B with a variable x, the function $\phi(x)(1 - x)^d$ has d roots on the unit circle i. e. $1 - x = 0$ when $x = 1$indicating that the process is non stationary hence the need for differencing.

NB: When $\phi(B)$ and $\theta(B)$ are both equal to unity (so that p and q are both zero) and d equals one, the model reduces to an ARIMA(0, 1, 0) model given by $X_t - X_{t-1} = Z_t$

Which is similar to a random walk model.

ARIMA models focus on how to predict the conditional mean of future values based on current and past data.

3.7 Models for changing variance

While ARIMA models focus on how to predict the conditional means of future values based on current and past data, time series models for heteroscedasticity can predict the variability of future values based on current and past data. ARCH and GARCH models can be used in such a case. The ARCH (1) model is the simplest GARCH model and is similar to an AR (1) model.

Most financial time series are often not normally distributed which means that prediction intervals may not be symmetric. Models for changing variance are rarely applied directly to the observed data. The derived series $\{Y_t\}$ should be (approximately) uncorrelated but may have a variance that changes through time. Then it may be represented in the form

$$Y_t = \sqrt{\sigma_t^2}\,\epsilon_t \tag{3.7.1}$$

where ϵ_t denotes a sequence of independent random variables with zero mean and unit variance and σ_t (which isn't observable directly), may be thought of as the local conditional standard deviation of the process.

3.7.1 ARCH models

An ARCH model is formally a regression model with the conditional volatility as the response variable and the past lags of the squared returns as the covariates. Let y_t be a random variable drawn from the conditional density function $f(y_t|y_{t-1})$. Engle (1982) proposed the following model:

$$y_t = \epsilon_t \sqrt{\sigma_t^2}, \qquad\qquad \{\epsilon_t\} \approx iid \quad N(0,1) \tag{3.7.2}$$

where σ_t^2 is the (positive) function of $\{y_s, s < t\}$ defined by

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2$$

with $\alpha_0 > 0$ and $\alpha_i \geq 0, i = 1, \cdots, q$ being known parameters.
σ_t^2 is the conditional variance of y_t given $\{y_s, s < t\}$

The ARCH model of order q, ARCH (q), assumes that σ_t^2 is linearly dependent on the last q squared values of the time series. Thus Y_t is said to follow an ARCH (1) model if it satisfies equation (3.7.1) where the conditional variance evolves through the ordinary equation:

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 \tag{3.7.3}$$

Where the constant parameters α_0 and α_1 are chosen to ensure that σ_t must be non negative and y_{t-1} denotes the observed value of the derived series at time $(t-1)$. Note the absence of an 'error' term in (3.7.3)

The simplest is the ARCH (1) process whereby

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2$$

ϵ_t is a sequence of identically and independently distributed random variables each with zero mean and unit variance and is independent of $y_{t-j}, j = 1, 2, \cdots$. Adding the assumption of normality, it can be more directly expressed in terms of ψ, the information set available at time t. Using conditional densities, $y_t|\psi_{t-1} \approx N(0, \sigma_t^2)$

$$\sigma_{t|t-1}^2 = \alpha_0 + \alpha_1 y_{t-1}^2 \tag{3.7.4}$$

The innovation ϵ_t is presumed to have unit variance so that the conditional variance of y_t equals $\sigma_{t|t-1}$ i.e.

$$E(y_t^2|y_{t-j}, j = 1, 2, \cdots) = E(\sigma_{t|t-1}^2 \epsilon_t^2|y_{t-j}, j = 1, 2, \cdots)$$

$$= \sigma_{t|t-1}^2 E(\epsilon_t^2|y_{t-j}, j = 1, 2, \cdots)$$

$$= \sigma^2_{t|t-1} E(\epsilon^2_t)$$

$$= \sigma^2_{t|t-1}$$

Problems with the ARCH (q) model

1. Non stationarity may be generated due to long lag in the conditional variance equation.

2. How does one decide on the value of q? For very many series, ARCH processes with fairly large orders are required to capture the dynamics in the conditional variances.

3. The non-negativity constraints may be violated resulting in forecasting of negative variances.

The GARCH model by Bollerslev (1986) can get round these problems by offering a more parsimonious model (it avoids overfitting)

3.7.2 GARCH models

In practice, q in the ARCH (q) model is often large. A more parsimonious representation is the GARCH model, GARCH (p,q) is often used to investigate volatility and persistence.Thus the ARCH model has been generalized to allow linear dependence of the conditional variance, σ^2_t, on the past values of σ^2_t as well as on past (squared) values of the series. The GARCH model of order (p, q) assumes the conditional variance depends on the squares of the last q values of the series and on the last p values of σ^2_t

A GARCH (p, q) process is given by:

$$\sigma^2_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i y^2_{t-i} + \sum_{j=1}^{p} \beta_j \sigma^2_{t-j} \tag{3.7.5}$$

where
$p \geq 0, \qquad q \geq 0, \qquad \alpha_0 > 0$
$\alpha_i \geq 0, for \quad i = 1, \cdots, q$
$\beta_j \geq 0, for \quad j = 1, \cdots, p$

For $p = 0$ the process reduces to the ARCH (q) process and for $p = q = 0$. ϵ_t is simply white noise.

Opening up equation (3.7.5) of the GARCH (p, q) model is as shown below:

$$\sigma^2_{t|t-1} = \alpha_0 + \beta_1 \sigma^2_{t-1|t-2} + \cdots + \beta_p \sigma^2_{t-p|t-p-1} + \alpha_1 y^2_{t-1} + \alpha_2 y^2_{t-2} + \cdots + \alpha_q y^2_{t-q} \tag{3.7.6}$$

In terms of the backward shift operator, B, the model can be expressed as:
$[1 - \beta_1 B - \cdots - \beta_p B^p] \sigma^2_{t|t-1} = \alpha_0 + (\alpha_1 B + \cdots + \alpha_q B^q) y^2_t$

In this paper, a simple GARCH (1,1) model is used:
Mean equation: $\qquad r_t = \mu + \epsilon_t$

The mean equation follows an ARMA model. μ gives the mean while ϵ_t is the error term we would like to model with GARCH. The error term ϵ_t in the ARMA equation is what is broken down into $\epsilon_t = \sigma_t z_t$ whereby $z_t \approx D(0, 1)$ and D is the conditional distribution of choice. σ_t is given by:
Variance equation: $\qquad \sigma^2_t = \alpha_0 + \alpha_1 \epsilon^2_{t-1} + \beta_1 \sigma^2_{t-1}$
where
$\alpha_0 \geq 0, \quad \alpha_1 \geq 0, \quad \beta_1 \geq 0$

r_t = Exchange rate return at time t

μ = average returns

ϵ_t =residual returns with $\epsilon_t = \sigma_t Z_t$ where Z_t is a random walk and σ_t^2 is the conditional variance

The parameters α_0, α_1,β_1 must satisfy $(\alpha + \beta) < 1$ for stationarity. (If $\alpha + \beta = 1$ then the process does not have finite variance, although it can be shown that the squared observations are stationary after taking first differences leading to Integrated GARCH.) GARCH model is thus an approximation to a higher order ARCH model.

Relaxing the Gaussian assumption in (3.7.2), better fits to the data can be obtained and we suppose instead that the distribution of Y_t given $\{Y_s, s < t\}$ has a heavier tailed zero-mean distribution such as student-t distribution. To incorporate such a distribution we can define a general GARCH (p, q) process as a stationary process $\{Y_t\}$ satisfying (3.7.5) and the generalized form of, (3.7.2) i. e.

$$Y_t = \epsilon_t \sigma_t, \{\epsilon_t\} \approx iidN(0,1)$$

For modelling purposes, it is assumed in addition that either $\epsilon_t \approx N(0,1)$ as in (3.7.2) or that $\sqrt{\frac{v}{v-2}}t_v \approx t_v, v > 2$

where

t_v denotes the student-t distribution with v degrees of freedom

EGARCH Model

EGARCH (Nelson, 1991)models are used to model leverage effect i.e. the fact that volatility tends to increase more after a negative return than a positive one. The model is given by:

$$ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{q} \alpha_i g(\epsilon_{t-i}) + \sum_{j=1}^{p} \beta_j ln(\sigma_{t-i}^2)$$

where

$$g(\epsilon_t) = \alpha\epsilon_t + \gamma\{|\epsilon_t|\} - E\{|\epsilon_t|\}$$

$ln(\sigma_t^2)$ can be negative i.e. there is no constraint in the parameters. To understand g, note that

$g(\epsilon_t) = -\gamma E(|\epsilon_t|) + (\gamma + \alpha)|\epsilon_t| \qquad\qquad \epsilon_t > 0$

and

$g(\epsilon_t) = -\gamma E(|\epsilon_t|) + (\gamma - \alpha)|\epsilon_t| \qquad\qquad \epsilon_t < 0$

typically, $-1 < \hat{\alpha} < 0$ so that $0 < \gamma + \alpha < \gamma - \alpha$

$$E[|\epsilon_{t-1}|] = \begin{cases} \sqrt{\frac{2}{\pi}}, & \text{for a standard normal variable} \\ \sqrt{\frac{v-2}{2}}.\Gamma(\frac{v-1}{2})\Gamma\frac{v}{2}, & \text{for student-t} \\ \lambda.2^{\frac{1}{v}}\frac{\Gamma(\frac{2}{v})}{\Gamma(\frac{1}{v})}, & \text{for GED} \end{cases}$$

where v is the degrees of freedom and $\lambda = [2^{\frac{-2}{v}}\Gamma(\frac{1}{v})\Gamma(\frac{3}{v})]^{\frac{1}{2}}$

$\alpha_0, \alpha_1, \beta_1$and γ_1 are the parameters to be estimated. α_1 gives the magnitude effect, β_1 the persistence so that when it is high, volatility takes a long time to die out. γ is a measure of the leverage effect i.e. if

1. $\gamma = 0$, then the model is symmetric

2. $\gamma < 0$, then negative shocks generate more volatility than positive shocks

3. $\gamma > 0$, then positive shocks generate more volatility than negative shocks

The EGARCH(1,1) model which is used in this paper is given by:

$$ln(\sigma_t^2) = \alpha_0 + \beta_1 ln(\sigma_{t-1}^2) + \alpha_1 \left[|\frac{\epsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}}| - E(\epsilon_{t-1}) \right] + \gamma \frac{\epsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}}$$

Mixed ARMA-GARCH Models

ARMA models are conditional expectation models while GARCH models are conditional variance models. Autocorrelation in the return series indicates that an ARMA model might be appropriate while autocorrelation in the squared series indicates that a GARCH model may be appropriate thus if there is autocorrelation in the series as well as in the squared series, a combination of ARMA and GARCH models is appropriate. The process, say X_t is defined as:

$a(B)X_t = c + \beta(B)\epsilon_t$

where

a(B) is a polynomial of degree p in B

$\beta(B)$ is a polynomial of degree q in B

ϵ_t is a GARCH process

3.8 Conditional Distributions of the error ϵ_t

The GARCH (p,q) model is usually specified with normal innovations Z_t distributed as standard normal:
$Z_t \approx N(0,1)$

which then implies estimation of the GARCH (p,q) parameters using maximum likelihood estimates yet in most cases the distribution of the innovations exhibit fatter tails than that of a normal distribution. Evidence thus shows that the conditional distribution of ϵ_t is by extension, non-normal. In this paper, four different error distributions out of the six listed below are considered to see which best captures the distribution of the innovations:

1. Normal Distribution
   $f(z_t) = \frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2}Z_t^2), -\infty < z < \infty$

2. Skewed Normal Distribution

   $f(z) = \frac{1}{\omega\pi} \exp \frac{-(z-\xi)^2}{2\omega^2} \int_{-\infty}^{\alpha\frac{z-\xi}{\omega}} exp\frac{-t^2}{2}dt, -\infty < z < \infty$

3. Student-t Distribution
   Bollerslev (1986) proposed the standardized student-t distribution with $\upsilon > 2$ degrees of freedom. The student-t distribution as given by:

   $f(z, \upsilon) = \frac{\Gamma\frac{(\upsilon+1)}{2})}{\Gamma\frac{\upsilon}{2}\sqrt{\pi(\upsilon-2)}}(1 + \frac{z_t^2}{\upsilon-2})^{-\frac{(\upsilon+1)}{2}}$

   where $\Gamma(.)$ is the gamma function. The distribution is symmetric about zero with $\upsilon > 2$. At $\upsilon > 4$, its kurtosis becomes $3(\upsilon - 2)(\upsilon - 4)$ which is larger than that of a normal distribution i. e. 3. As $\upsilon \to \infty$, the distribution converges to a standard normal.

4. Skewed Student-t Distribution

This was proposed by Fernandez and Steel(1998). Lambert and Laurent (2001) then applied this to GARCH . It is given by:

$$f(z; \mu, \sigma, \upsilon, \lambda) = \begin{cases} qr(1 + \frac{1}{\upsilon-2}(\frac{q(\frac{z-\mu}{\sigma})+p}{1-\lambda})2)^{-\frac{\upsilon+1}{2}}, & \text{if } z < -\frac{a}{b}, \\ qr(1 + \frac{1}{\upsilon-2}(\frac{q(\frac{z-\mu}{\sigma})+p}{1+\lambda})2)^{-\frac{\upsilon+1}{2}}, & \text{if } z \geq -\frac{p}{q} \end{cases}$$

with $-1 < \lambda < 1$. The constants p, q and r are given below:

$p = 4\lambda(\frac{\upsilon-2}{\upsilon-1})$

$q = 1 + 3\lambda - p^2$

$r = \frac{\Gamma(\frac{\upsilon+1}{2})}{\sqrt{\pi(\upsilon-2)}\Gamma(\frac{\upsilon}{2})}$

where μ and σ^2 are the mean and variance of the skewed student-t distribution respectively.

In this paper, the normal distribution and the students-t distribution performances are compared using the AIC and log-likelihood values. Their variants, the skew-normal and skew-student-t are also considered.

3.9 Maximum Likelihood Estimates

The estimation of a GARCH (p,q) model with t-distribution follows a Quasi Maximum Likelihood Estimation (QMLE) since the normality assumption is violated. The likelihood function of a GARCH model can be derived for the case of normal innovations. Given parameters ω, α and β, the conditional variance can be computed recursively by the formula:

$\sigma^2_{t|t-1} = \omega + \alpha r^2_{t-1} + \beta \sigma^2_{t-1|t-2}$ for $t \geq 2$,

with the initial value , $\sigma^2_{1|0}$, set under the stationarity assumption as the stationary unconditional variance

$\sigma^2 = \frac{\omega}{1-\alpha-\beta}$.

The conditional pdf given below is used:

$f(r_t|r_{t-1}, \cdots, r_1) = \frac{1}{\sqrt{2\pi\sigma^2_{t|t-1}}} exp[\frac{-r^2_t}{(2\sigma^2_{t|t-1})}]$

and the joint pdf:

$f(r_n, \cdots, r_1) = f(r_{n-1}, \cdots, r_1)f(r_n|r_{n-1}, \cdots, r_1)$

Iteration of the joint pdf formula and taking logs gives the log-likelihood formula as shown below:

$L(\omega, \alpha, \beta) = -\frac{n}{2}log(2\pi) - \frac{1}{2}\sum_{i=1}^{n} \left\{ log(\sigma^2_{t-1|t-2}) + \frac{r^2_t}{\sigma^2_{t|t-1}} \right\}$

There is no closed form solution for the maximum likelihood estimators of ω, α,and β, but they can be computed by maximizing the log-likelihood function numerically.

Condsider the observed sequence Y_t denoted by:

$$Y_t = C_0 + \epsilon_{0t}, t = 1, \cdots, n,$$

It is assumed that (ϵ_{0t}) is a GARCH (1,1) process i.e.

$$\epsilon_{0t} = Z_t \sigma_{0t}, \quad F_t = \sigma(\epsilon_{0s}, s \leq t),$$
with z_t being a sequence of iid random variables and

$$\sigma_{0t}^2 = \omega_0(1 - \beta_0) + \alpha_0 \epsilon_{0t-1}^2 + \beta_0 \sigma_{0t-1}^2 \text{ a.s}$$
The strict stationary solution of the latter equation is given by:

$$\sigma_{0t}^2 = \omega_0 + \alpha_0 \sum_{k=0}^{\infty} \beta_0^k \epsilon_{0t-1-k}^2 \text{ a.s}$$

if $E[ln(\beta_0 + \alpha_0 Z^2)] < 0$ holds. The vector of the parameters describing the process is thus given by:

$$\theta_0 = (C_0, \omega_0, \alpha_0, \beta_0).$$

The model for the unknown parameters $\theta = (C, \omega, \alpha, \beta)'$ is given by:

$$Y_t = C + \epsilon_t, t = 1, \cdots, n,$$

and

$$\sigma_t^2(\theta) = \omega(1 - \beta) + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2(\theta), t = 2, \cdots, n$$

and with the initial condition $\sigma_1^2(\theta) = \omega$.

The process for the conditional variance is thus given as:
$$\sigma_t^2 = \omega + \alpha \sum_{k=0}^{t-2} \beta^k \epsilon_{t-1-k}^2$$

Then define the compact space Θ by:

$$\Theta = \theta : C_t \leq C \leq C_d, 0 \leq \omega_l \leq \omega \leq \omega_d, 0 \leq \alpha_l \leq \alpha_d,$$

$$0 \leq \beta_l \leq \beta \leq \beta_d < 1$$

$$\subseteq \{\theta : E[\ln(\beta + \alpha Z^2)] < 0\}$$

A further assumption made is that $\theta_0 \in \Theta$ and $\alpha_0 > 0, \beta_0 > 0$. GARCH(1,1) processes usually assume that (Z_t) are iid random variables such that $Z_t \approx N(0,1)$. Assuming the likelihood function is Gaussian, the log-likelihood function is of the form:

$$L_t(\theta) = \frac{1}{2T} \sum_{t=1}^{T} l_t(\theta), \qquad \text{where } l_t(0) = -(\ln \sigma_t^2(\theta) + \frac{\epsilon_t^2}{\sigma_t^2(\theta)})$$

Since the likelihood function does not need to be Gaussian, L_T is called the quasi-likelihood function. The log likelihood function of the student-t distribution is given by:

$$l_t = N\left\{log\Gamma(\frac{\upsilon+1}{2}) - log\Gamma(\frac{\upsilon}{2}) - log[\pi(\upsilon - 2)]\right\} - \frac{1}{2}\sum_{n=1}^{N}\left[log(\sigma_t^2) + (1+\upsilon)log(1 + \frac{z_t^2}{\upsilon-2})\right]$$

where υ is the degrees of freedom, $2 < \upsilon \leq \infty$ and $\Gamma(.)$ is the gamma function.

The log likelihood function of the standardized skewed student-t distribution is given by:

$$l_t = N \left\{ log\Gamma(\tfrac{\upsilon+1}{2}) - log\Gamma(\tfrac{\upsilon}{2}) - 0.5 log[\pi(\upsilon-2)] + log(\tfrac{2}{\xi+\frac{1}{\xi}}) + log(s) \right\}$$

$$- 0.5 \sum_{n=1}^{N} \left\{ log\sigma_t^2 + (1+\upsilon)log\left[1 + \tfrac{(sz_t+m)^2}{\upsilon-2}\xi^{-2I_t} \right] \right\}$$

where:

$I_t = 1$ if $z_t \geq -\tfrac{m}{s}$ and   -1 if $z_t < -\tfrac{m}{s}$

ξ is the asymmetry parameter, υ is the degree of freedom of the distribution.

$$m = \frac{\Gamma(\frac{\upsilon+1}{2})\sqrt{(\upsilon-2)}}{\sqrt{\pi}\Gamma(\frac{\upsilon}{2})}\left(\xi - \tfrac{1}{\xi}\right)$$

and

$$s = \sqrt{\left(\xi^2 + \tfrac{1}{\xi^2} - 1\right) - m^2}$$

3.10   Model selection criteria

In this paper, the Akaike's(1973) Information Criterion is used for model selection. The criterion says to select the model that minimizes:

AIC $= -2log$(maximum likelihood) + 2k

where $k = p + q + 1$ if the model contains an intercept or a constant term and $k = p + q$ otherwise. The addition of the $2(p+q+1)$ or $2(p+q)$ serves as a "penalty function " thus ensuring the selection of a parsimonious model. k is the number of parameters in the model. The value of k yielding the minimum AIC specifies the best model. The lower the AIC value, the better the model fit.

AIC balances the error of the fit against the number of parameters.

3.11   Stationarity

A stationary time series is one in which the behavior of the process does not change over time. A process $\{y_t\}$ is said to be strictly stationary if the joint distribution of $y_{t_1}, y_{t_2}, \ldots, y_{t_n}$ is the same as that of $y_{t_1-k}, y_{t_2-k}, \ldots, y_{t_n-k}$ for all choices of time points $t_1, t_2, \ldots, t_n$ and all choices of time lag k. Thus when $n = 1$ the (univariate) distribution of y_t is the same as that of y_{t-k} for all t and k. It follows that $E[y_t] = E[y_{t-k}]$ for all t and k so that the mean function is constant for all time. Also $Var(y_t) = Var(y_{t-k})$ for all t and k so that the variance is also constant over time. A stochastic process $\{y_t\}$ is said to be weakly/second order stationary if:

1. The mean function is constant over time

2. Variance is constant

3. The autocovariance doesn't depend on time

In this paper, the ADF test has been used to test for stationarity.

3.11.1   Testing for unit roots

In practice it is usually difficult to distinguish between a stationary and non-stationary process. For ARIMA models, it is usually difficult to answer the question of whether the parameter 'd' is exactly equal to one. If it is, then a unit root is present i. e. the original data is non-stationary but the first differences are stationary.

Other than observing the linear decay of the ACF of a time series as a sign that the underlying time series is non stationary and requires differencing, it is useful to quantify the evidence of non-stationarity in the data generating mechanism and this can be done via hypothesis testing. Consider the model

$$Y_t = \alpha Y_{t-1} + X_t \ \text{ for } t = 1, 2, \cdots$$

where $\{X_t\}$ is a stationary process. The process $\{Y_t\}$ is non-stationary if the coefficient $\alpha = 1$, but it is stationary if $|\alpha| < 1$. Suppose that $\{X_t\}$ is an AR(k) process: $X_t = \phi_1 X_{t-1} + \cdots + \phi_k X_{t-k} + \epsilon_t$. Under the null hypothesis that $\alpha = 1$, $X_t = Y_t - Y_{t-1}$. Letting $a = \alpha - 1$, we have

$$Y_t - Y_{t-1} = (\alpha - 1)Y_{t-1} + X_t$$
$$= aY_{t-1} + \phi_1 X_{t-1} + \cdots + \phi_k X_{t-k} + \epsilon_t$$
$$= aY_{t-1} + \phi_1(Y_{t-1} - Y_{t-2}) + \cdots + \phi_k(Y_{t-k} - Y_{t-k-1}) + \epsilon_t$$

where $a = 0$ under the hypothesis that Y_t is difference nonstationary. On the other hand, if $\{Y_t\}$ is stationary so that $-1 < \alpha < 1$, then it can be verified that Y_t still satisfies an equation similar to the one above but with different coefficients

The ADF-test is thus used to test for unit roots. The null hypothesis states that there is a unit root, i.e. $\alpha = 1$

3.12   Skewness

The skewness of a random variable, say Y, is defined by $\frac{E(Y-\mu)^3}{\sigma^3}$ where μ and σ are the mean and standard deviation of Y, respectively. It can be estimated by the sample skewness

$$g_1 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^3}{(n\hat{\sigma}^3)}$$

where $\hat{\sigma}^2 = \frac{\sum (Y_i - \bar{Y})^2}{n}$ is the sample variance.

The thickness of the tail of a distribution relative to that of a normal distribution is often measured by the (excess) kurtosis defined as:

$$\frac{E(Y-\mu)^4}{\sigma^4 - 3}$$

For the normal distributions, kurtosis is always equal to zero. A distribution with positive kurtosis is called a heavy-tailed distribution whereas it is called light-tailed if its kurtosis is negative. The kurtosis can be estimated by sample kurtosis.

$$g_2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^4}{(n\hat{\sigma}^4 - 3)}$$

Another test for normality is the Jarque-Bera test based on the fact that a normal distribution has zero skewness and zero kurtosis. Assuming identically and independent data, $Y_1, Y_2, \cdots, Y_n$, the Jarque-Bera statistic is defined as :

$$JB = \frac{ng_1^2}{6} + \frac{ng_2^2}{24}$$

where g_1 is sample skewness and g_2 is sample kurtosis.

Chapter 4

DATA ANALYSIS AND RESULTS

4.1 DATA

Daily exchange rate data is used to model volatility of exchange rates of the USD/KES. The data span from 23rd October 1993 to 21st March 2014, a total of 7456 observations. The data was obtained from www.oanda.com.

The variable modelled is the first difference of the natural logarithm of the exchange rate as shown by the equation below:

$$r_t = \log(\frac{X_t}{X_{t-1}})$$

where
r_t = Daily exchange rate return
X_t = Current day exchange rate
X_{t-1} = Previous day exchange rate

4.2 EXPLORATORY DATA ANALYSIS

4.2.1 EXCHANGE RATES DATA

The time plot of the exchange rates data is as shown in figure 4.1:

KES/USD Exchange Rates
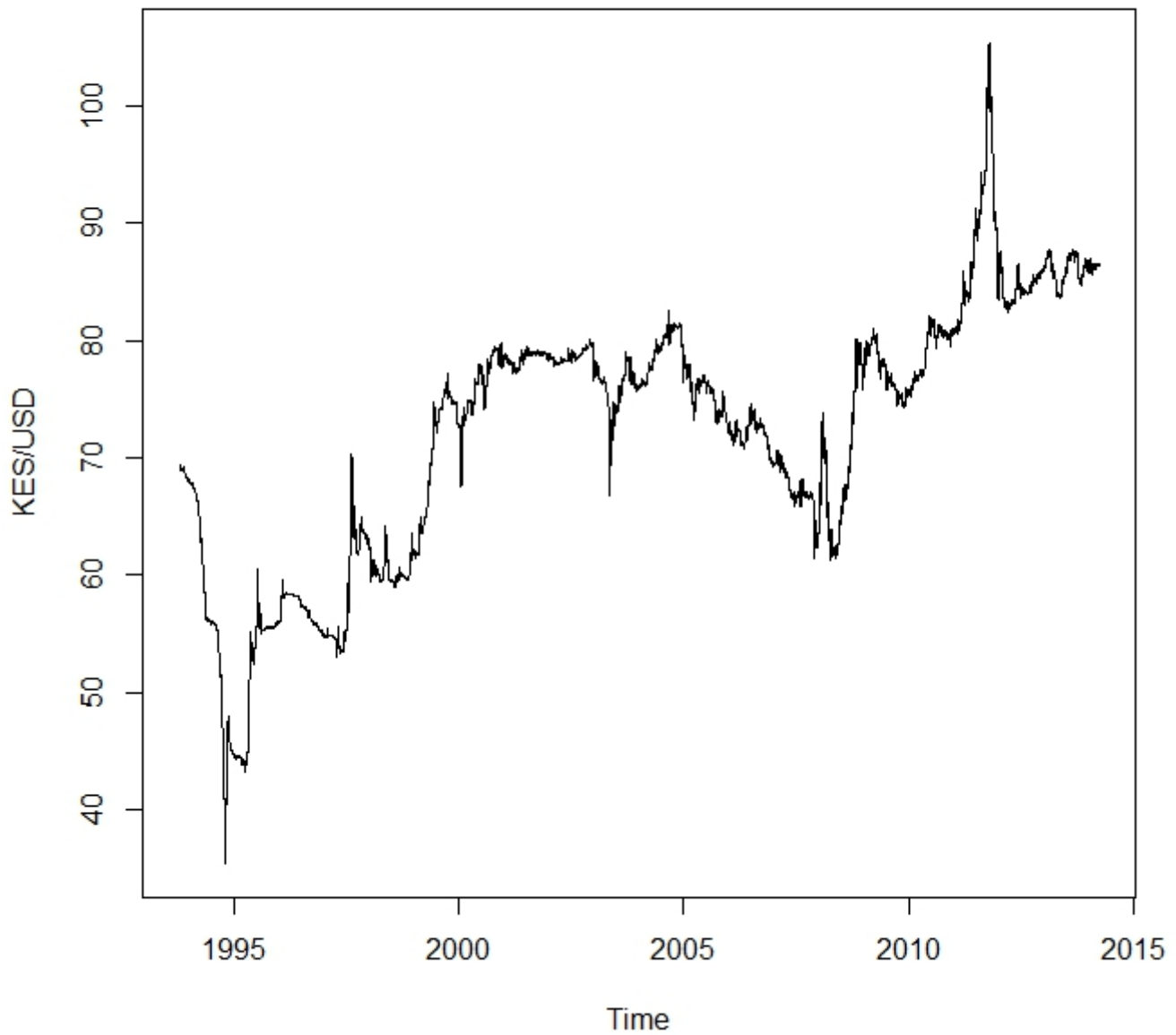
Figure 4.1: Time plot of USD/KES exchange rates data

Further tests carried out i.e. plots of the histogram fig.4.2 and Q-Q plot fig.4.3 of the exchange rates time series data show that the data is not normal but is heavy tailed.

Histogram of exchange rates data
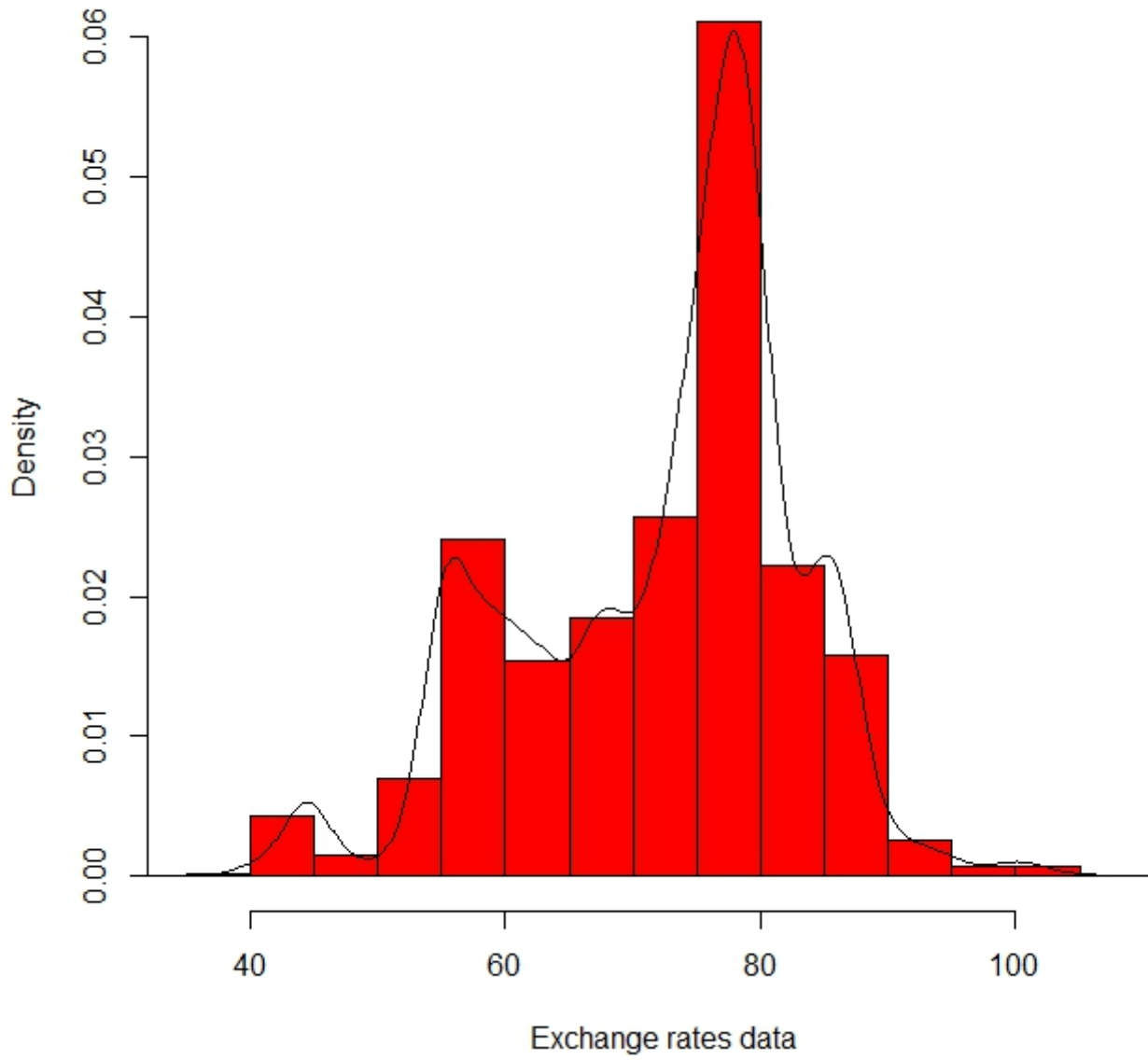
Figure 4.2: Histogram of KES/USD exchange rates data

Normal Q-Q Plot
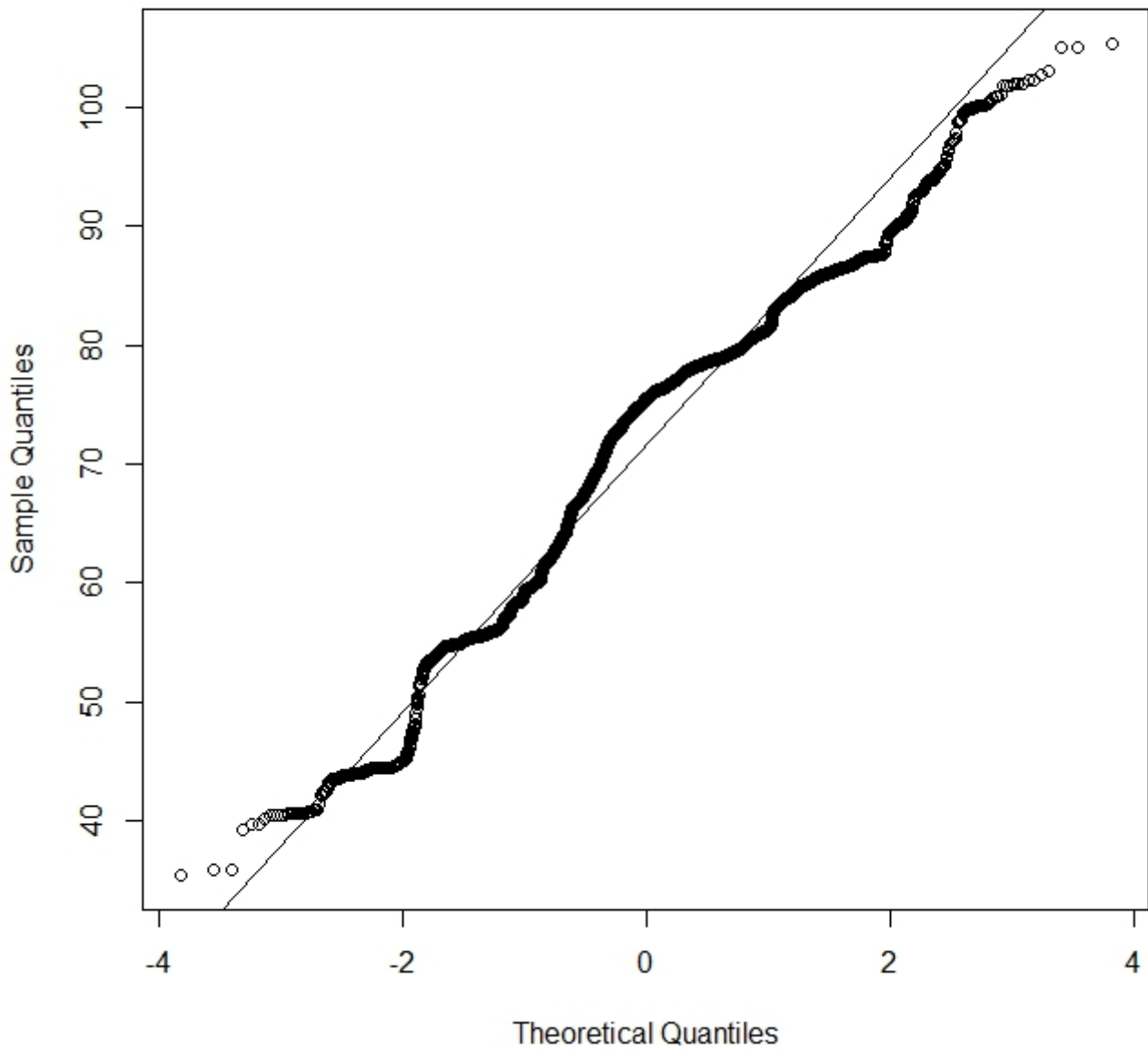
Figure 4.3: Q-Q plot of KES/USD exchange rates data

From figure 4.1 it is observed that the series is not stationary. This is also evidenced in the corellogram of the original time series which does not come down to zero indicating that the series is not stationary.
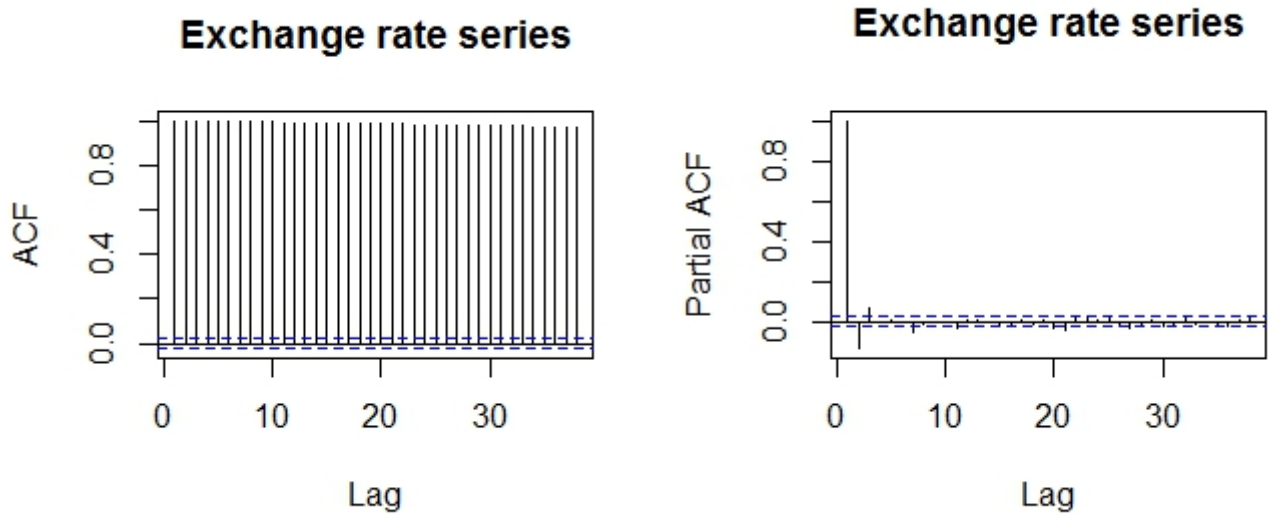
Figure 4.4: ACF and PACF of original series

Another test for stationarity is the ADF unit root test. The results are as follows:

Augmented Dickey-Fuller Test
data: DERtimeS
Dickey-Fuller = -2.6503, Lag order = 1, p-value = 0.3029
alternative hypothesis: stationary

The p value of 0.3029 leads us to accept the null hypothesis at 1%, 5% and even 10% level of significance and conclude that the series is non stationary.

4.2.2 Testing for ARCH effects

Various ARMA models we fitted to the data and the one with the lowest AIC and highest LLH selected. On the resulting residuals of the fit, ARCH effects were tested. The results were as follows:

Table 4.1: Comparison of AIC and log likelihood of Various ARMA models

	AIC	Log likelihood
ARMA(1,0)	-58384.88	29194.44
ARMA(0,1)	-58413.93	29208.97
ARMA(1,1)	-58417.44	29211.72
ARMA(1,2)	-58419.32	29213.66
ARMA(2,1)	-58421.54	29214.77
ARMA(2,2)	-58419.76	29214.88
ARMA(2,0)	-58423.54	29214.77
ARMA(0,2)	-58419.09	29212.54

Using the residuals of the ARMA(2,0) model, we then test for ARCH effects. The results are as follows:

ARCH LM-test; **Null hypothesis**: no ARCH effects
data: resid20
Chi-squared = 524.4517, df = 12, p-value ¡ 2.2e-16

Because of the small p-value, we reject the null hypothesis and conclude that there are ARCH effects, thereby justifying the use of GARCH models.

Table 4.2: Descriptive statistics of the KES/USD exchange rate

Min	1st Qu.	Median	Mean	3rd Qu.	Max
35.40	64.00	75.46	72.25	79.17	105.40

There is need to transform the data before further analysis can be carried out. Figure 4.5 and figure 4.6 plots show the logged series and the differenced series respectively. Logging the series attempts to stabilize the variance while differencing imposes stationarity as seen in the plot of the original series vs the logged and differenced series figure 4.8.

Log of Exchange rate series
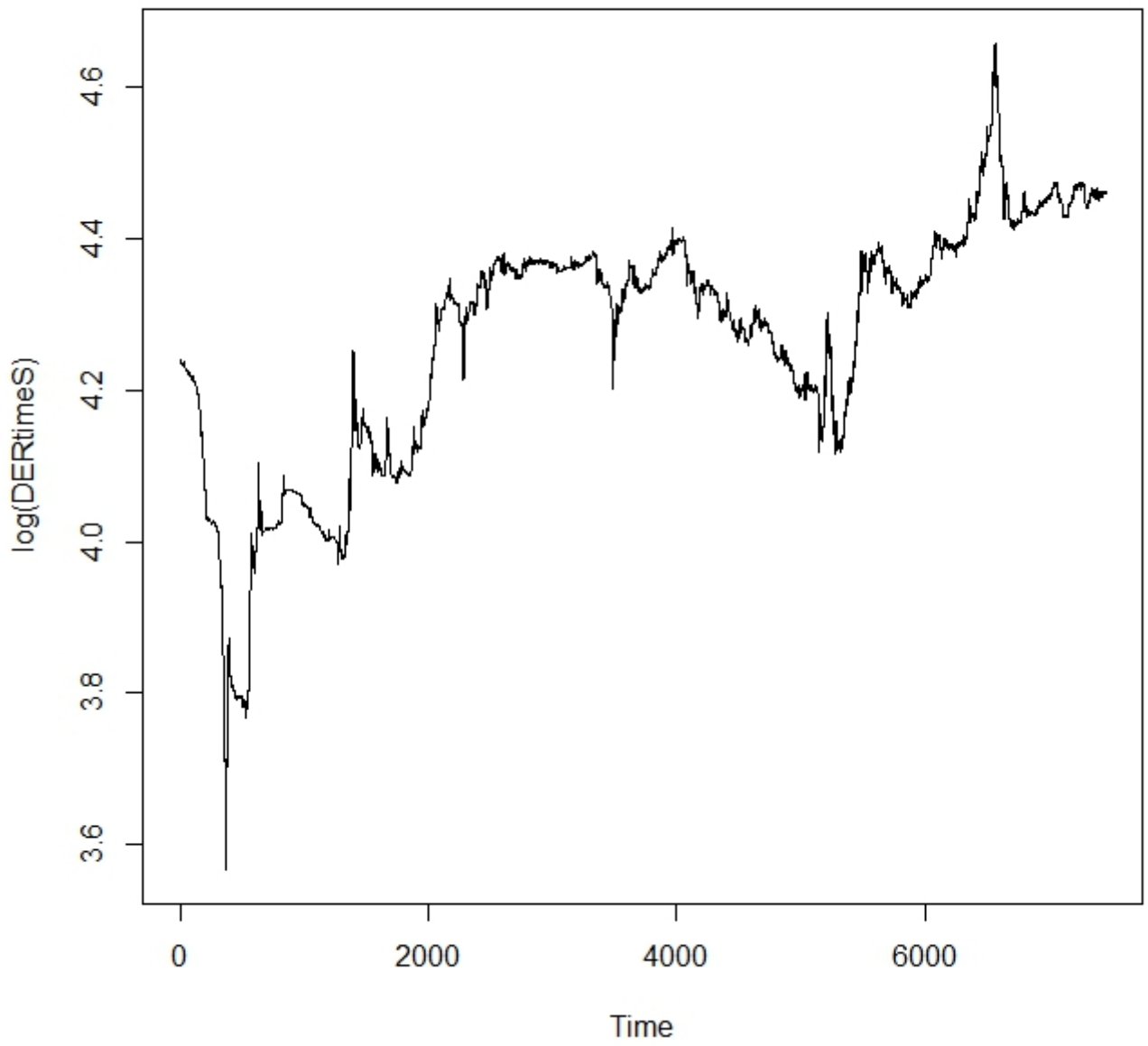
Figure 4.5: Plot of log of Exchange rate series

Plot of Difference of Exchange rate series
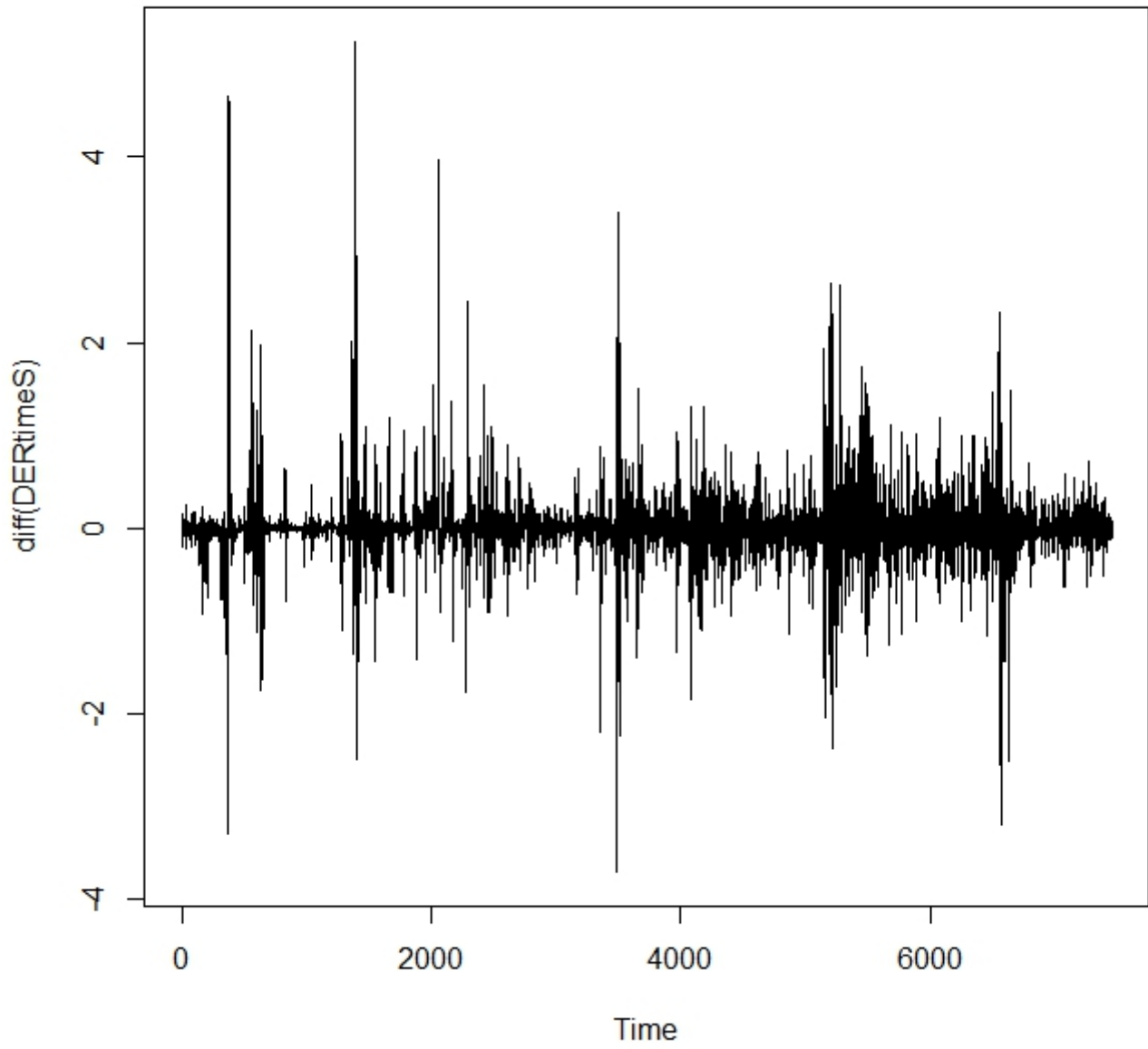
Figure 4.6: Plot of difference of Exchange rate series

Figure 4.7: Plot of logged and differenced Exchange rate series
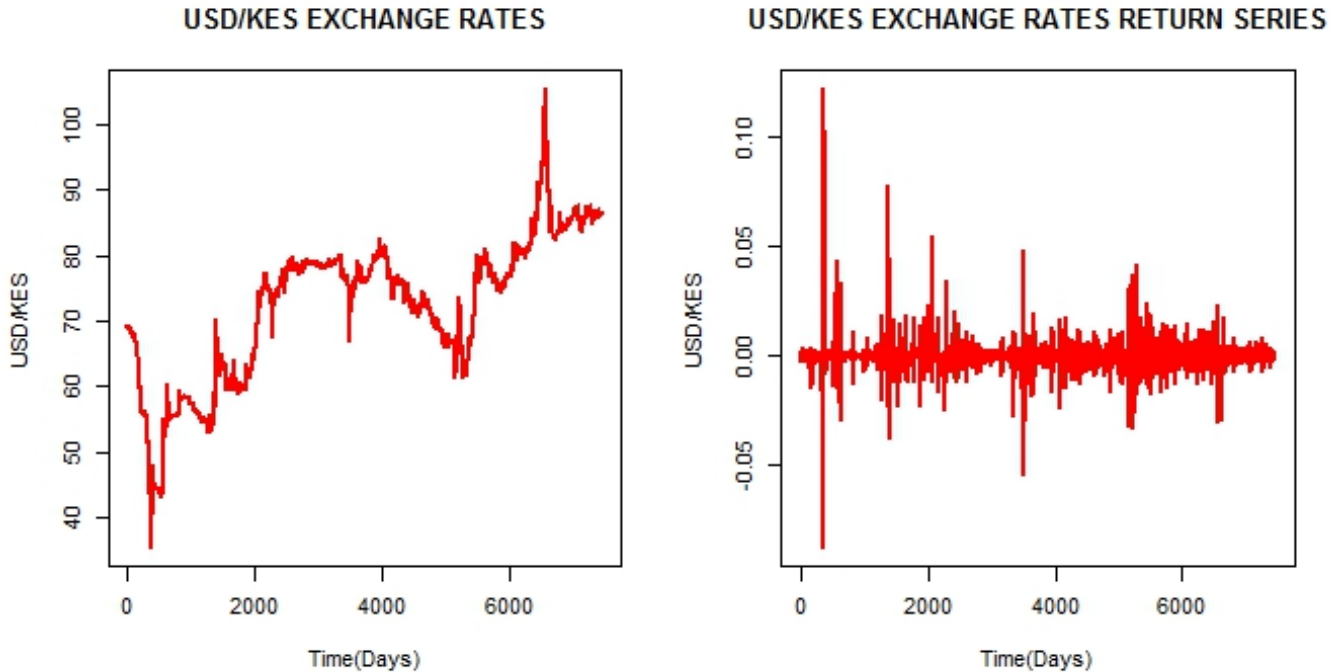
Figure 4.8: Plot of exchange rate series vs return series

From the figure 4.8, volatility clustering can be observed and this gives a hint that the data may not be independent and identically distributed, otherwise the variance would be constant over time.

4.2.3 Test for Stationarity

Three tests for stationarity were carried out on the transformed data:

ADF test results for the transformed series (logged and differenced series) are as follows:

data: NTER1
Dickey-Fuller = -59.5586, Lag order = 1, p-value = 0.01
alternative hypothesis: stationary

The ADF test of the transformed series gives a p-value smaller than 0.01 which leads us to reject the null hypothesis and conclude that the series is now stationary.

The PP-test gave the following results:

Phillips-Perron Unit Root Test Dickey-Fuller Z(alpha) = -6051.733, Truncation lag parameter = 11, p-value = 0.01
alternative hypothesis: stationary

The p value of 0.01 leads us to reject the null hypothesis and conclude that the exchange rate series is stationary. The KPSS test as well shows that the data is stationary. The results were as follows:

43

KPSS Level = 0.0815, **Truncation lag parameter** = 19, **p-value** = 0.1

4.2.4   EXCHANGE RATES RETURNS DATA

Descriptive statistics

Table 4.3: Descriptive statistics of the USD/KES exchange rate return series

Mean	Min.	Max	std.D	Skewness	Kurtosis	JB	ADF
2.973e-05	-8.782e-02	1.220e-01	0.004896149	3.503521	115.3206	4145648	-59.5586

From Table 4.3, the skewness and kurtosis tell us that the data is not normal. Data from a normal distribution tends to have a skewness of 0 and kurtosis of 3. The positive skewness tells us that the data is skewed to the right while the positive kurtosis tells us that the data is heavy tailed. The JB test statistic which tests for normality also confirms that the null hypothesis of normality should be rejected.

The ADF test for stationarity as seen in table 4.3 rejects the null hypothesis of a unit root, thus implying that the exchange rate return series data is now stationary.

A Q-Q plot, figure 4.9 of the exchange rate returns series, still indicates that the data is not normal.
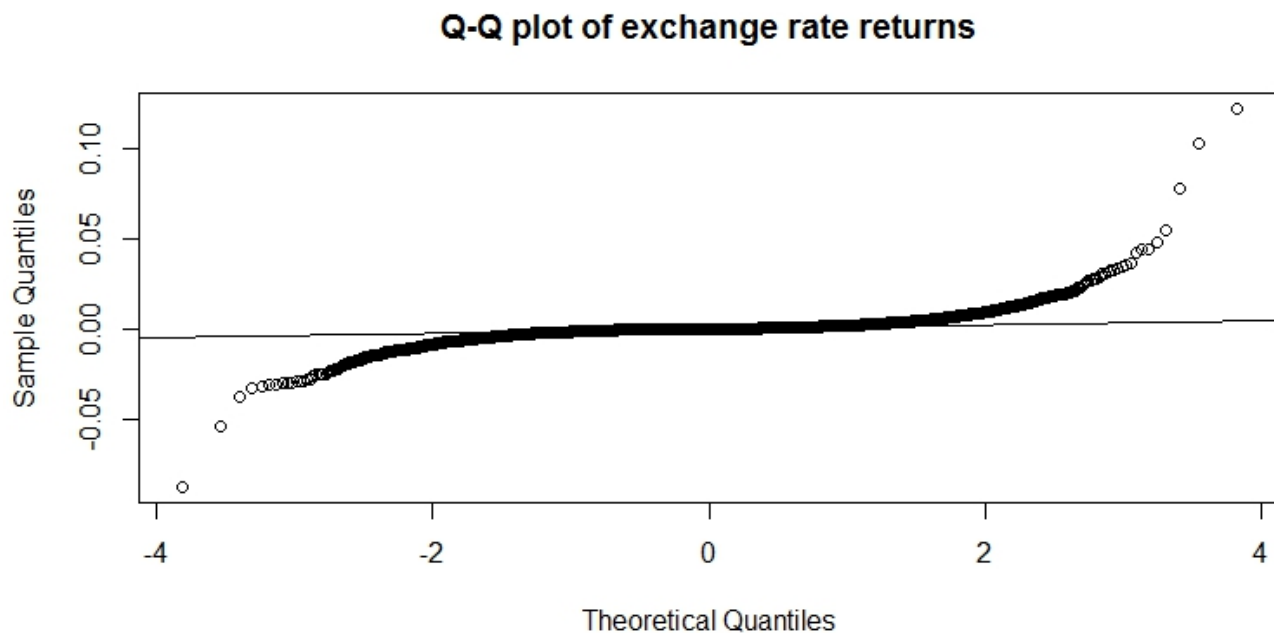
Figure 4.9: Q-Q plot of transformed data

Testing for ARCH effects

The ARCH LM-test is used to test for the presence of ARCH effects in the transformed data and the results obtained are as follows:

ARCH LM-test; Null hypothesis: no ARCH effects
data: NTER1
Chi-squared $= 545.34$, df $= 12$, p-value $< 2.2e - 16$

The p-value is small, which leads one to reject the null hypothesis and conclude that there are ARCH effects. This presence of ARCH effects justifies the use of GARCH models.
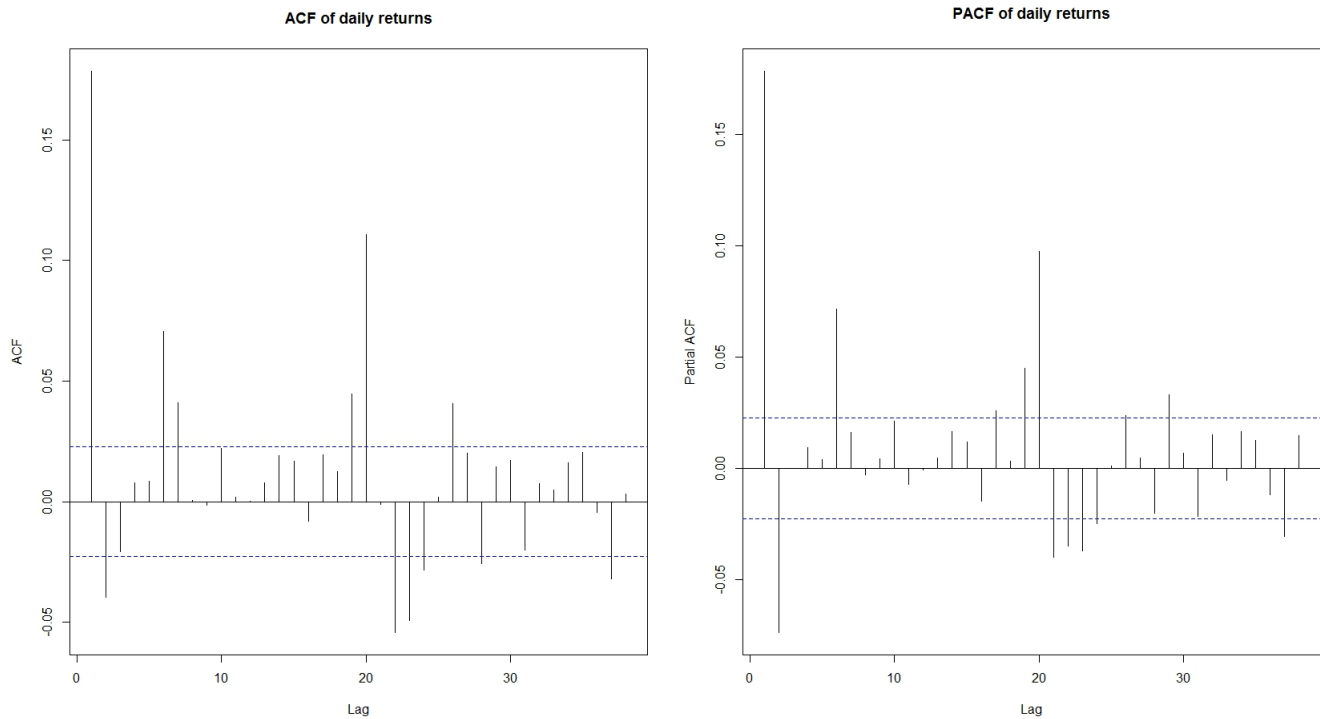
Figure 4.10: Sample ACF and PACF of daily returns
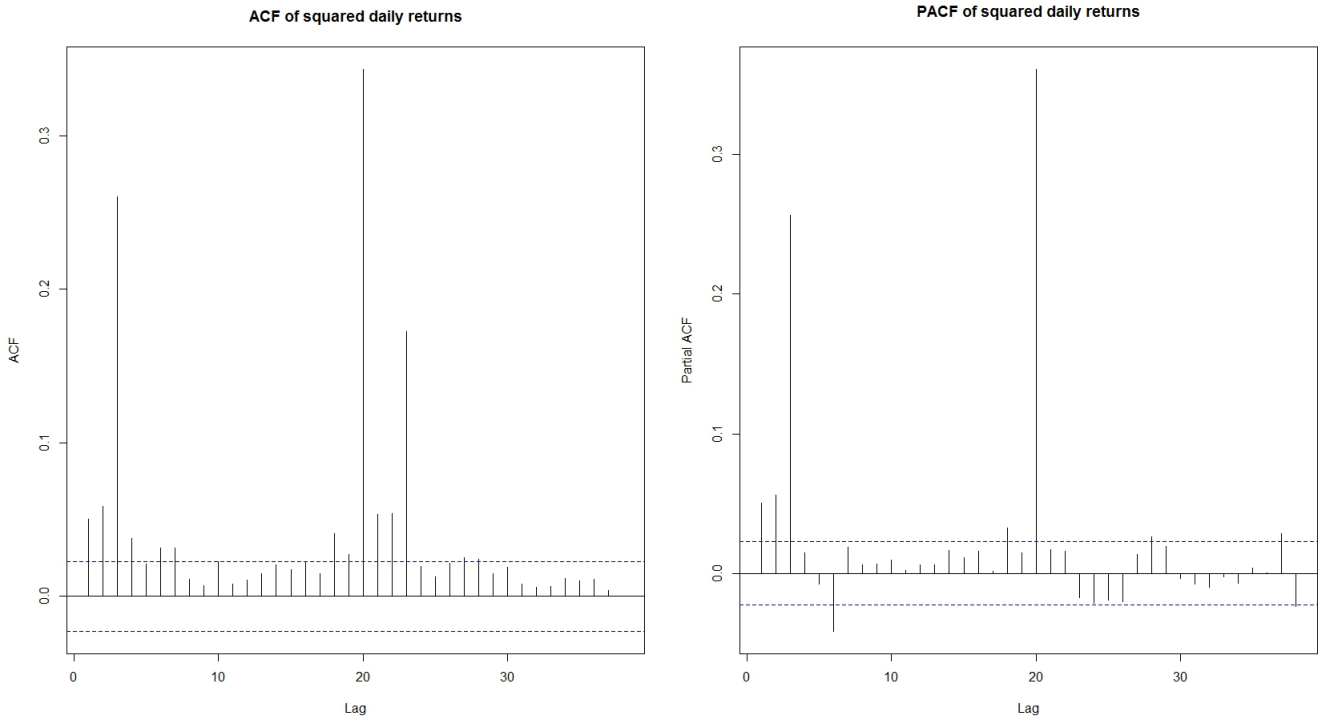
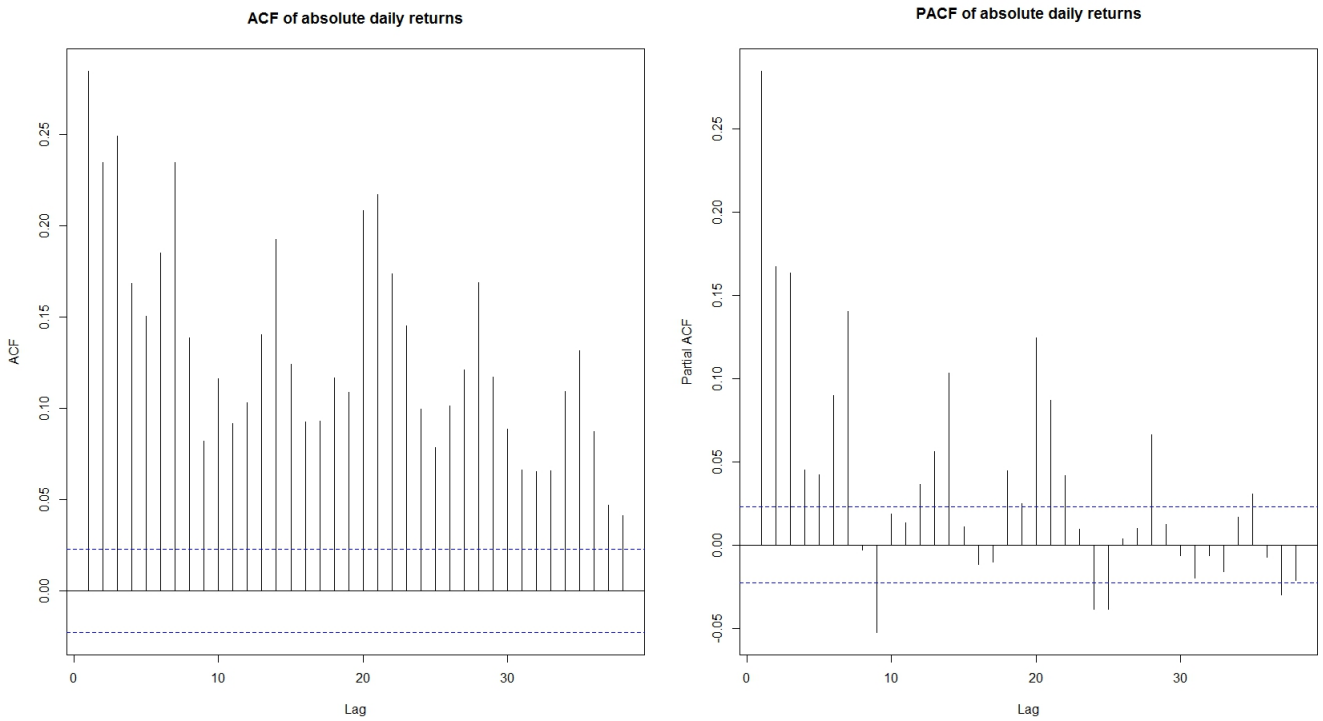Figure 4.11: Sample ACF and PACF of the squared daily returns
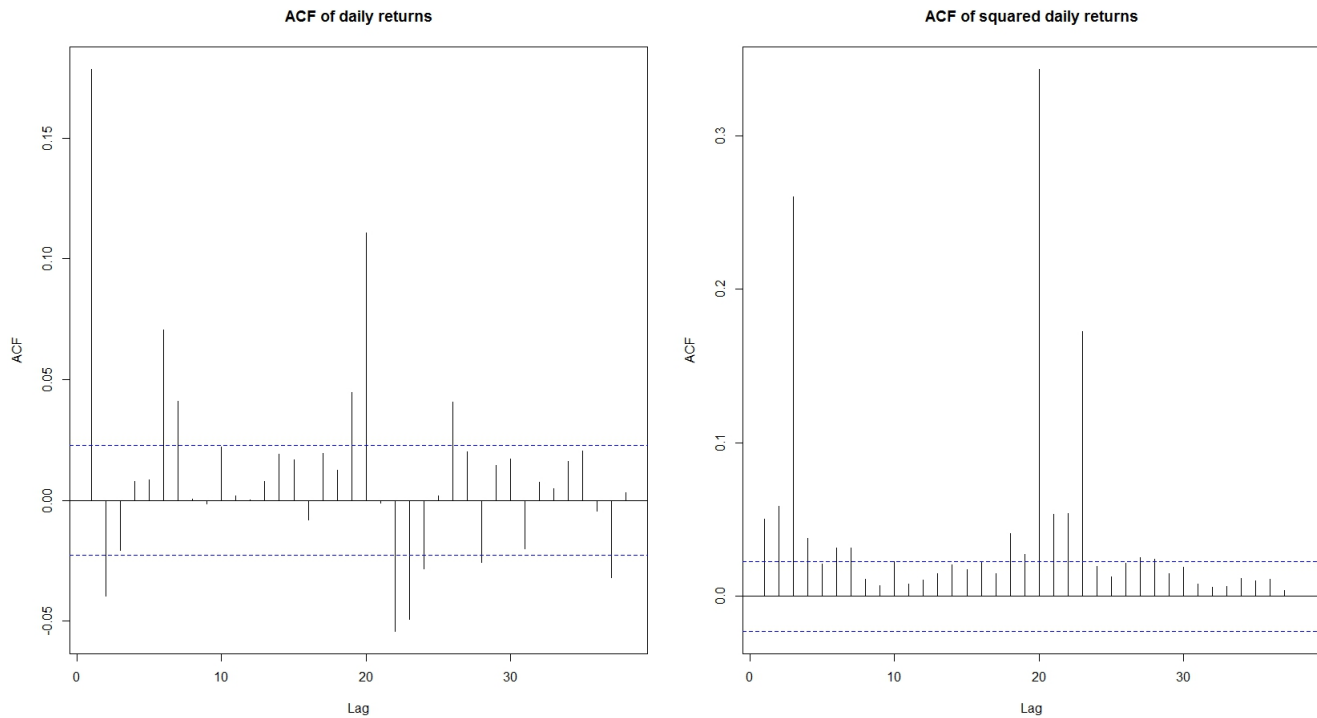

Figure 4.12: Sample ACF and PACF of the absolute daily returns

Figure 4.13: Sample ACF of daily returns and ACF of the squared daily returns

Test for Autocorrelation

The plots of the ACF of the squared returns, figure 4.11 and absolute returns, figure 4.12 display significant autocorrelations which further provide some evidence that the daily exchange rate returns are not independently and identically distributed.

The ACF plot of the KES/USD exchange rate series shows serial correlation and the PACF plot of the squared KES/USD exchange rate series indicates long term dependence.

Thus the assumptions underlying a GARCH model: serial correlation, presence of ARCH effects and long term dependence are satisfied.

4.2.5 The Model

Tests for autocorrelation in the exchange rate series indicate that there is autocorrelation in the series as well as in the squared series (Figure 4.13) thus a combination of an ARMA and GARCH model were considered appropriate.

Various $\text{ARMA}(p_A, q_A)/\text{GARCH}(1,1)$ as well as $\text{ARMA}(p_A, q_A)/\text{EGARCH}(1,1)$ models were considered for modelling the KES/USD exchange rate return series and a model was selected based on the one that had the least AIC and the highest Loglikelihood.

Table 4.4: ARMA(p_A, q_A) / GARCH(1,1) model with different conditional distributions

	Normal	Skewed Normal	Student -t	Skewed Student-t
ARMA(1,0)/GARCH(1,1)				
AIC	-8.687827	-8.691969	-9.404614	-9.404354
Log likelihood	32384.53	32400.97	35057	35057.03
ARMA(0,1)/GARCH(1,1)				
AIC	-8.689553	-8.693669	-9.405408	-9.40515
Log likelihood	32390.96	32407.31	35059.96	35059.99
ARMA(1,1)/GARCH(1,1)				
AIC	-8.691422	-8.695547	-9.405715	-9.405418
Log likelihood	32398.93	32415.3	35062.1	35061.99
ARMA(2,0)/GARCH(1,1)				
AIC	-8.692563	-8.696741	-9.405512	-9.405257
Log likelihood	32403.18	32419.75	35061.34	35061.39
ARMA(0,2)/GARCH(1,1)				
AIC	-8.692289	-8.696445	-9.405616	-9.405358
Log likelihood	32402.16	32418.65	35061.73	35061.77
ARMA(1,2)/GARCH(1,1)				
AIC	-8.698726	-8.700693	-9.405361	-9.405105
Log likelihood	32427.15	32435.48	35061.78	35061.83
ARMA(2,1)/GARCH(1,1)				
AIC	-8.697261	-8.690188	-9.405361	-9.405105
Log likelihood	32427.15	32435.48	35061.78	35061.83
ARMA(2,2)/GARCH(1,1)				
AIC	-8.699273	-8.701256	**-9.409205**	-9.408993
Log likelihood	32430.19	32438.58	35077.11	35077.32

From Table 4.3 the ARMA(2,2)/GARCH(1,1) model with a student-t conditional distribution appears to be the best with the lowest AIC of -9.409205 and a log likelihood of 35077.11. From Table 4.4, the ARMA(2,2)/EGARCH(1,1) with the skewed student-t conditional distribution appears to be the best model with an AIC of -9.454082 and a log likelihood of 35246.36.

Comparing the two models, the ARMA(2,2)/EGARCH(1,1) with skewed student-t conditional distribution is the one selected because compared to the ARMA(2,2)/GARCH(1,1) with student-t conditional distribution, it has a lower AIC and a higher log likelihood as shown in the table below.

Table 4.6: Comparison of AIC and log likelihood of ARMA(2,2)/GARCH(1,1) and ARMA(2,2)/EGARCH(1,1)

	AIC	Log likelihood
ARMA(2,2)/GARCH(1,1) (student-t distribution)	-9.409205	35077.11
ARMA(2,2)/EGARCH(1,1) (skewed student-t distribution)	-9.454082	35246.36

The model selected is thus the ARMA(2,2)/EGARCH(1,1) with skewed student-t conditional distribution. The asymmetric EGARCH model is also a better model than the symmetric GARCH model because

Table 4.5: ARMA(p_A, q_A)/ EGARCH(1,1) model with different conditional distributions

	Normal	Skewed Normal	Student -t	Skewed Student-t
ARMA(1,0)/EGARCH(1,1)				
AIC	-8.713435	-8.718282	-9.443562	-9.448017
Log likelihood	32480.97	32500.04	35203.16	35220.7
ARMA(0,1)/EGARCH(1,1)				
AIC	-8.717933	-8.72006	-9.444296	-9.448693
Log likelihood	32497.74	32506.46	35205.89	35223.28
ARMA(1,1)/EGARCH(1,1)				
AIC	-8.706284	-8.703347	-9.444615	-9.448944
Log likelihood	32455.32	32445.37	35208.08	35225.21
ARMA(1,2)/EGARCH(1,1)				
AIC	-8.661896	-8.721835	-9.444362	-9.44873
Log likelihood	32290.89	32515.28	35208.14	35225.42
ARMA(2,0)/EGARCH(1,1)				
AIC	-8.707094	-8.705358	-9.4444538	-9.448913
Log likelihood	32458.34	32452.87	35207.79	35225.1
ARMA(0,2)/EGARCH(1,1)				
AIC	-8.708871	-8.723315	-9.444623	-9.448991
Log likelihood	32464.96	32519.8	35208.11	35225.39
ARMA(2,1)/EGARCH(1,1)				
AIC	-8.662286	-8.719181	-9.444368	-9.448735
Log likelihood	32292.34	32505.39	35208.15	35225.44
ARMA(2,2)/EGARCH(1,1)				
AIC	-8.717815	-8.709768	-9.449894	**-9.454082**
Log likelihood	32500.3	32471.3	35229.75	**35246.36**

unlike the GARCH model, it can account for leverage effects. Also, another significant advantage of the EGARCH model is that even if the parameters are negative, σ_t^2 will be positive because the variable modelled is $(\ln \sigma_t^2)$. The conditional mean equation with the error term following a conditional heteroscedastic process is given by:

Mean Equation:

$\phi(B)X_t = \theta(B)Z_t$ i.e.

$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$ (AR process)

using the backward shift operator, this becomes:

$X_t = \phi_1 B X_t + \phi_2 B^2 X_t + Z_t$

$X_t - \phi_1 B X_t - \phi_2 B^2 X_t = Z_t$

$X_t(1 - \phi_1 B - \phi_2 B^2) = Z_t$ $\qquad$ where $(1 - \phi_1 B - \phi_2 B^2) = \phi(B)$

therefore:

$\phi(B)X_t = Z_t$

Also:

$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$ (MA process)
using the backward shift operator, this becomes:

$$X_t = Z_t + \theta_1 B Z_t + \theta_2 B^2 Z_t$$

$$X_t = Z_t(1 + \theta_1 B + \theta_2 B^2) \qquad\qquad \text{where } (1 + \theta_1 B + \theta_2 B^2) = \theta(B)$$

therefore:
$$\theta(B)Z_t = X_t$$

Hence the mean equation is given by:

$$\phi(B)X_t = \theta(B)Z_t$$

i.e.
$$X_t(1 - \phi_1 B - \phi_2 B^2) = (1 + \theta_1 B + \theta_2 B^2)Z_t$$

4.2.6   SIMULATION

A simulation with the model ARMA(2,2)/EGARCH (1,1) appears to fit the structure of the original return series as depicted in Figure 4.14 and Figure 4.15 below:
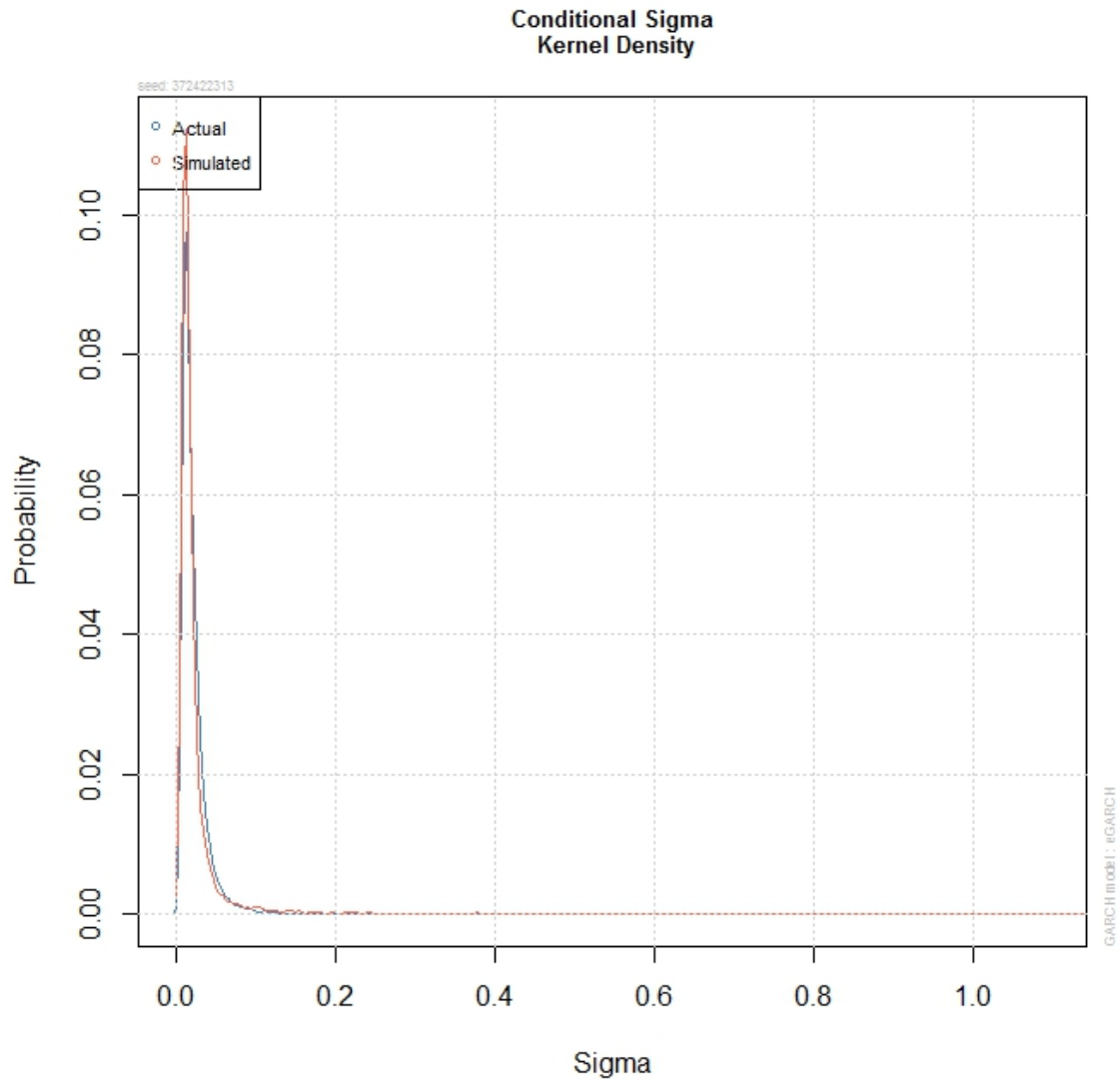
Figure 4.14: Conditional Standard Deviation Simulation
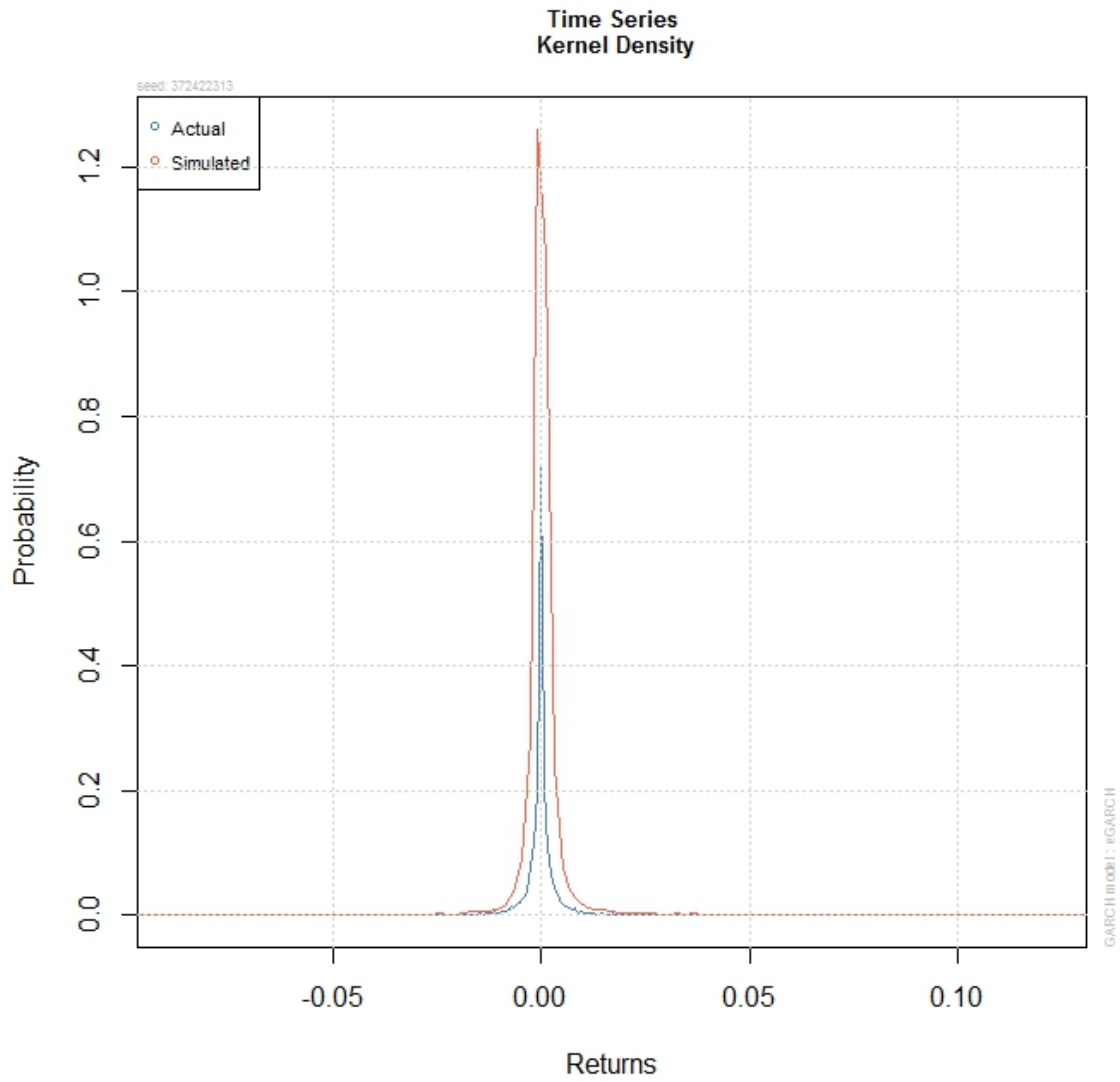
Figure 4.15: Return Series simulation Path Density
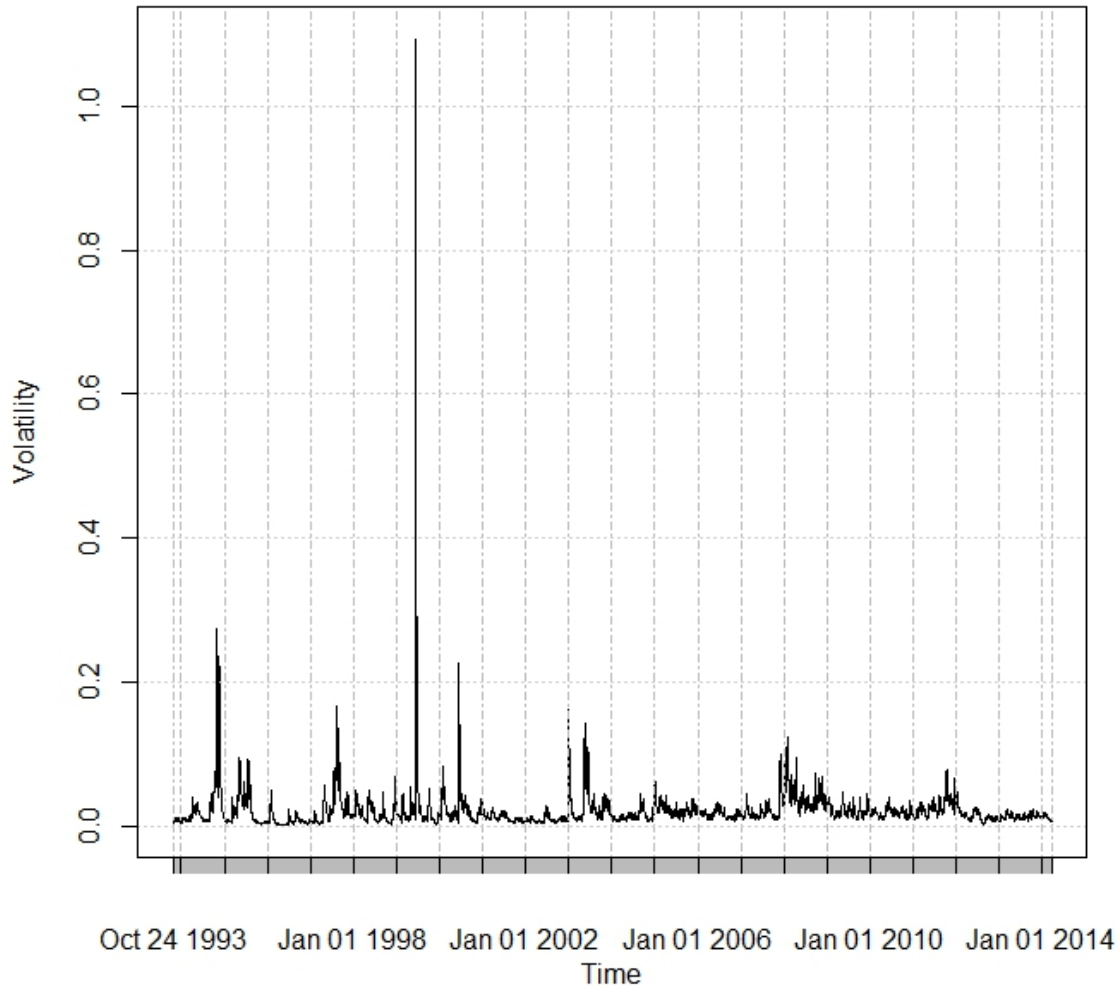
Figure 4.16: Model Volatility

Fig. 4.16 shows volatility clustering i.e. periods of calm and periods of high volatility. The image depicted is a model of volatility and not the true volatility which essentially is not observable. GARCH models are capable of modeling this volatility clustering but doesn't explain it.

The ARMA(2,2)/EGARCH(1,1) selected on the basis of AIC seems to be a good fit but for the sake of parsimony, we try and do simulations with other lower ARMA lags to see if their plots can provide similarly good fits, if not better. The plots for Conditional Standard Deviation Simulation and Return Series simulation Path Density are given below for comparison.

Conditional Sigma
Kernel Density
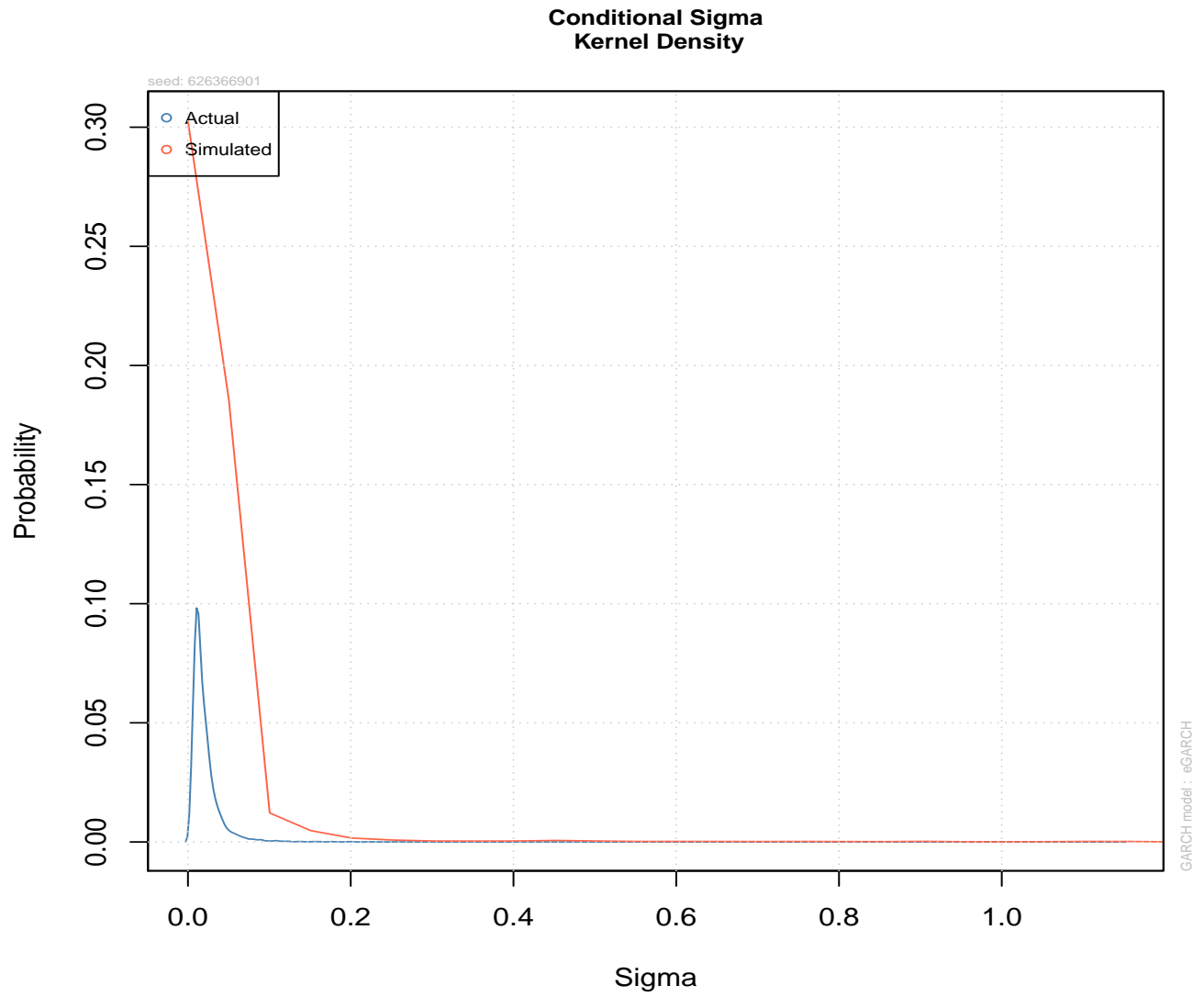
Figure 4.17: ARMA(2,1)/EGARCH (1,1) Conditional SD Simulation Density
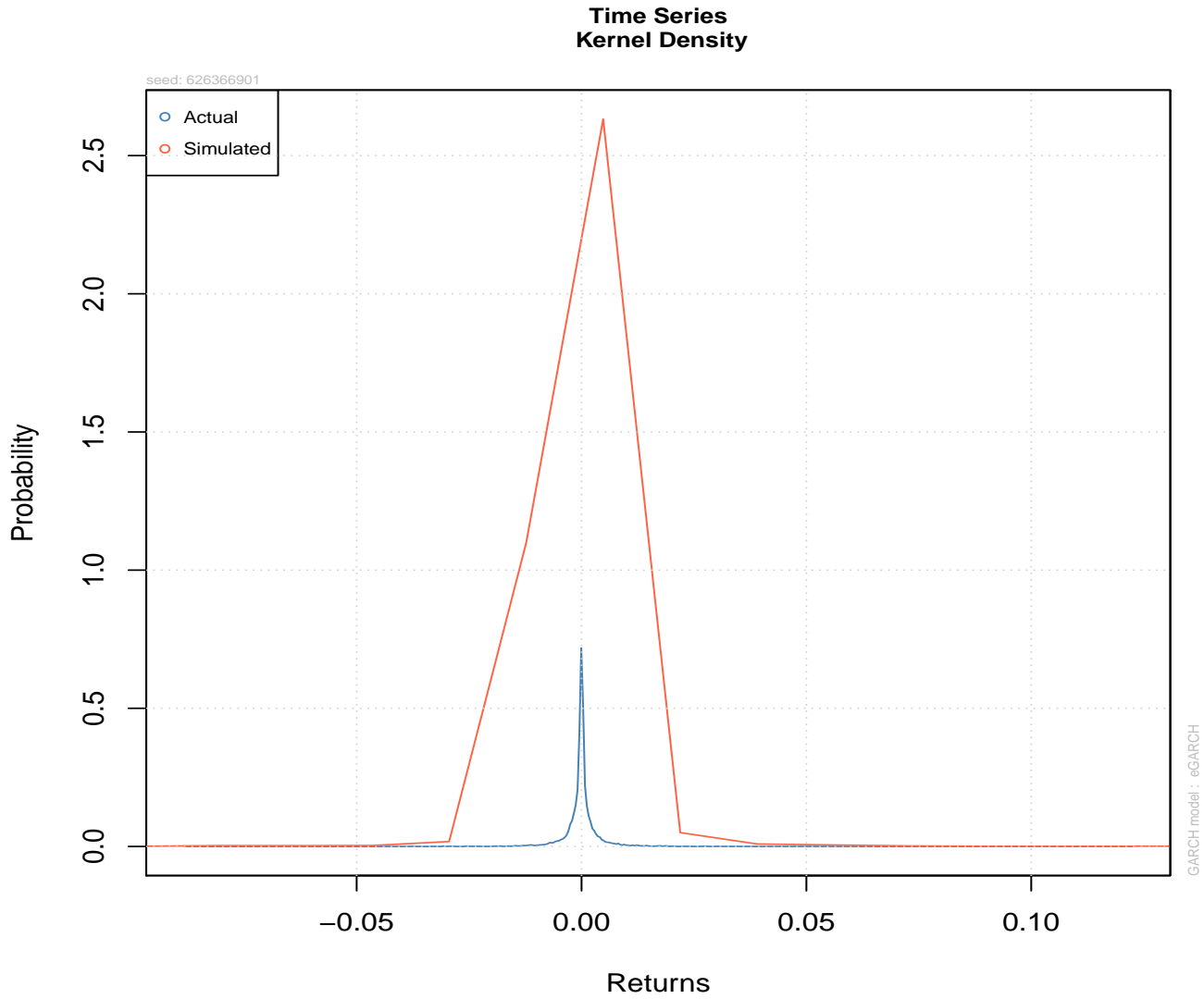
Figure 4.18: ARMA(2,1)/EGARCH (1,1) Return Series Simulation Path Density
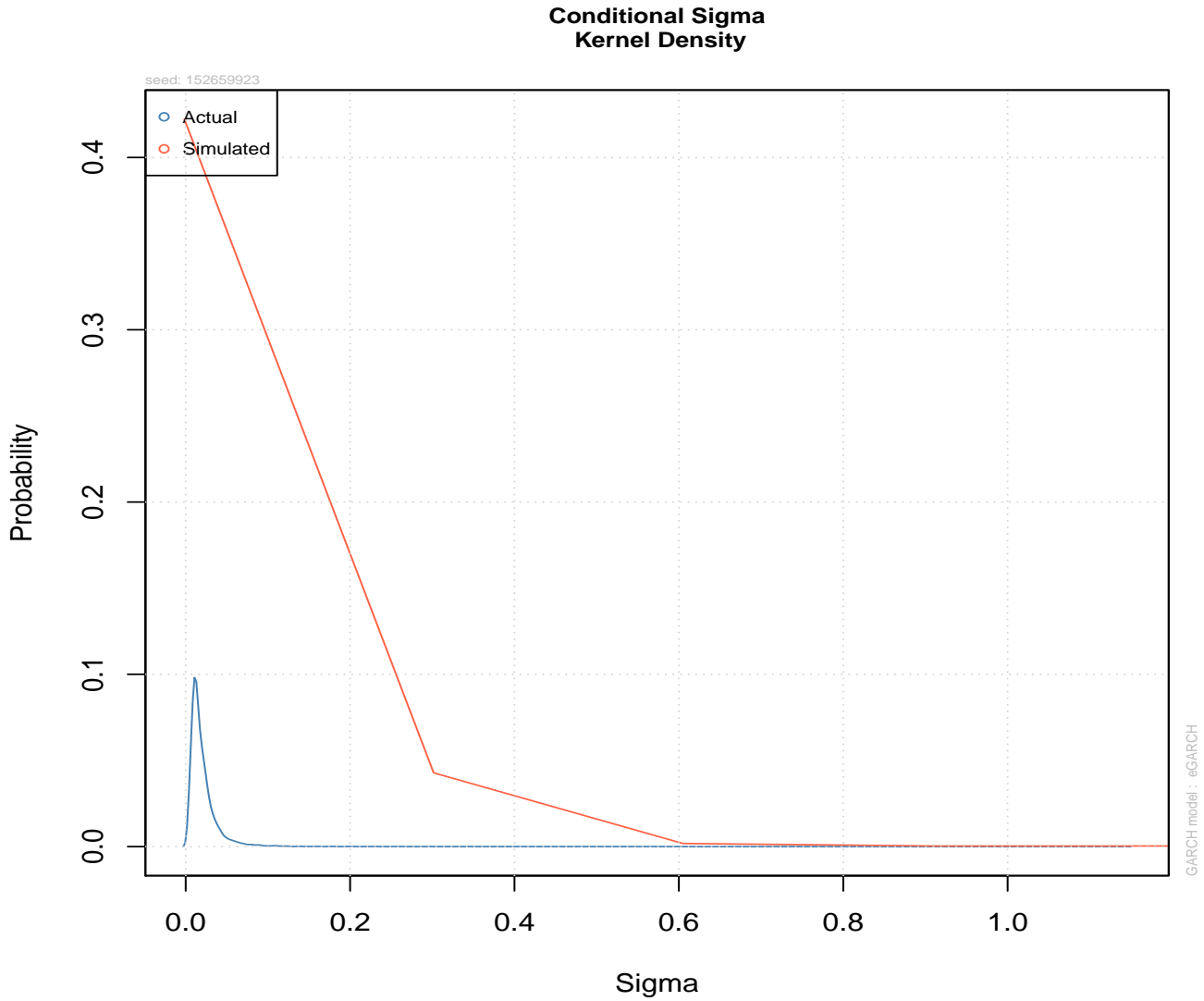
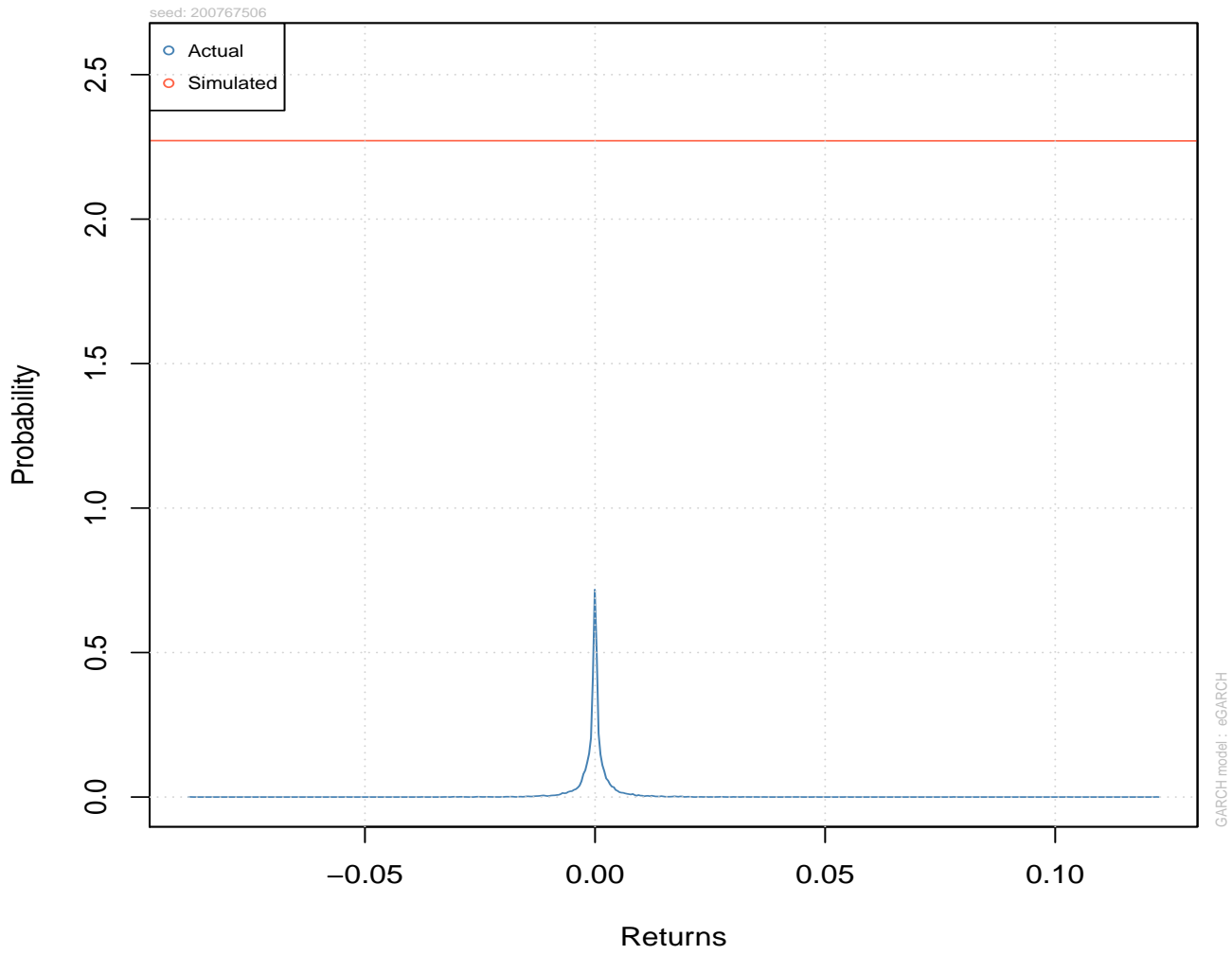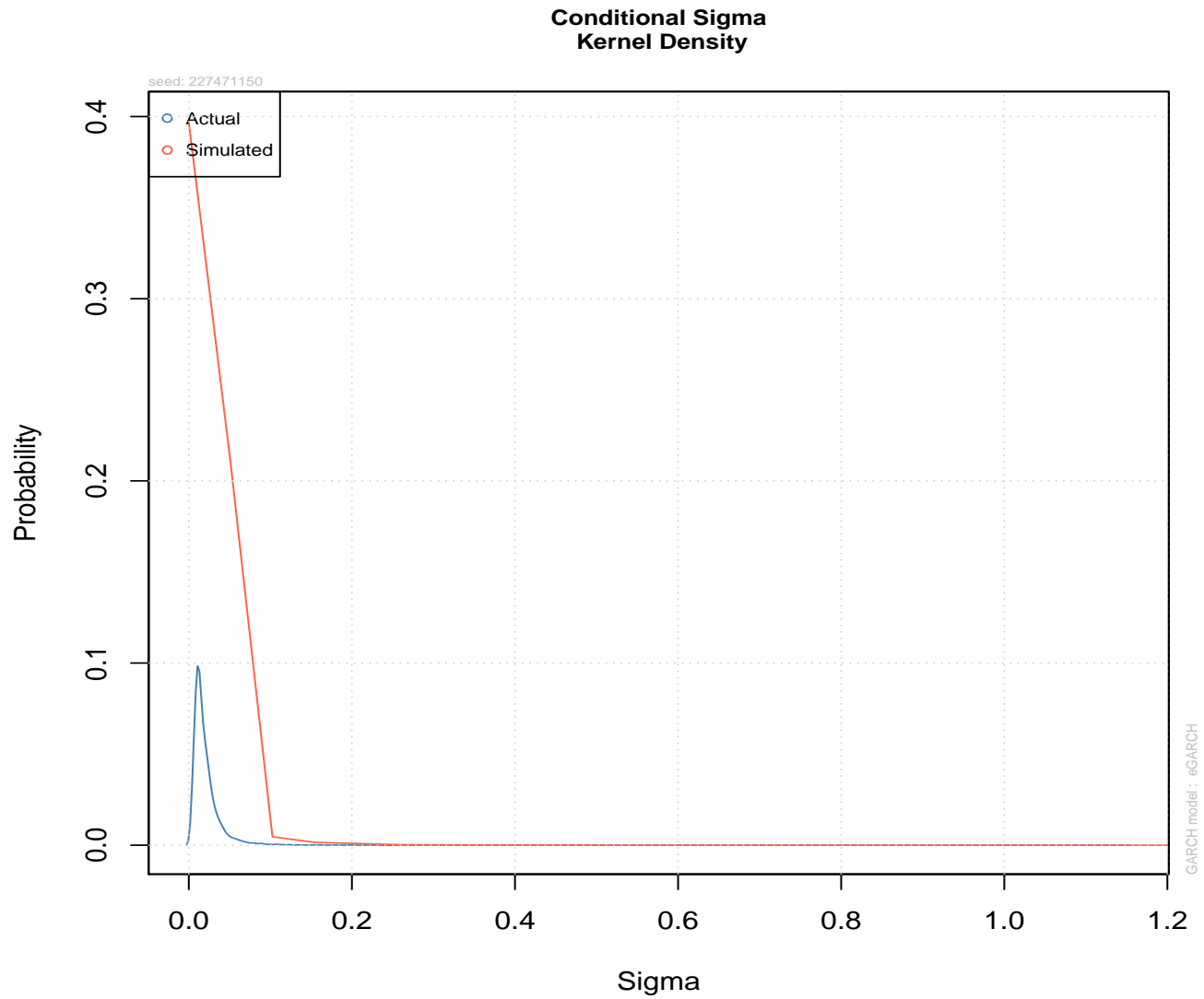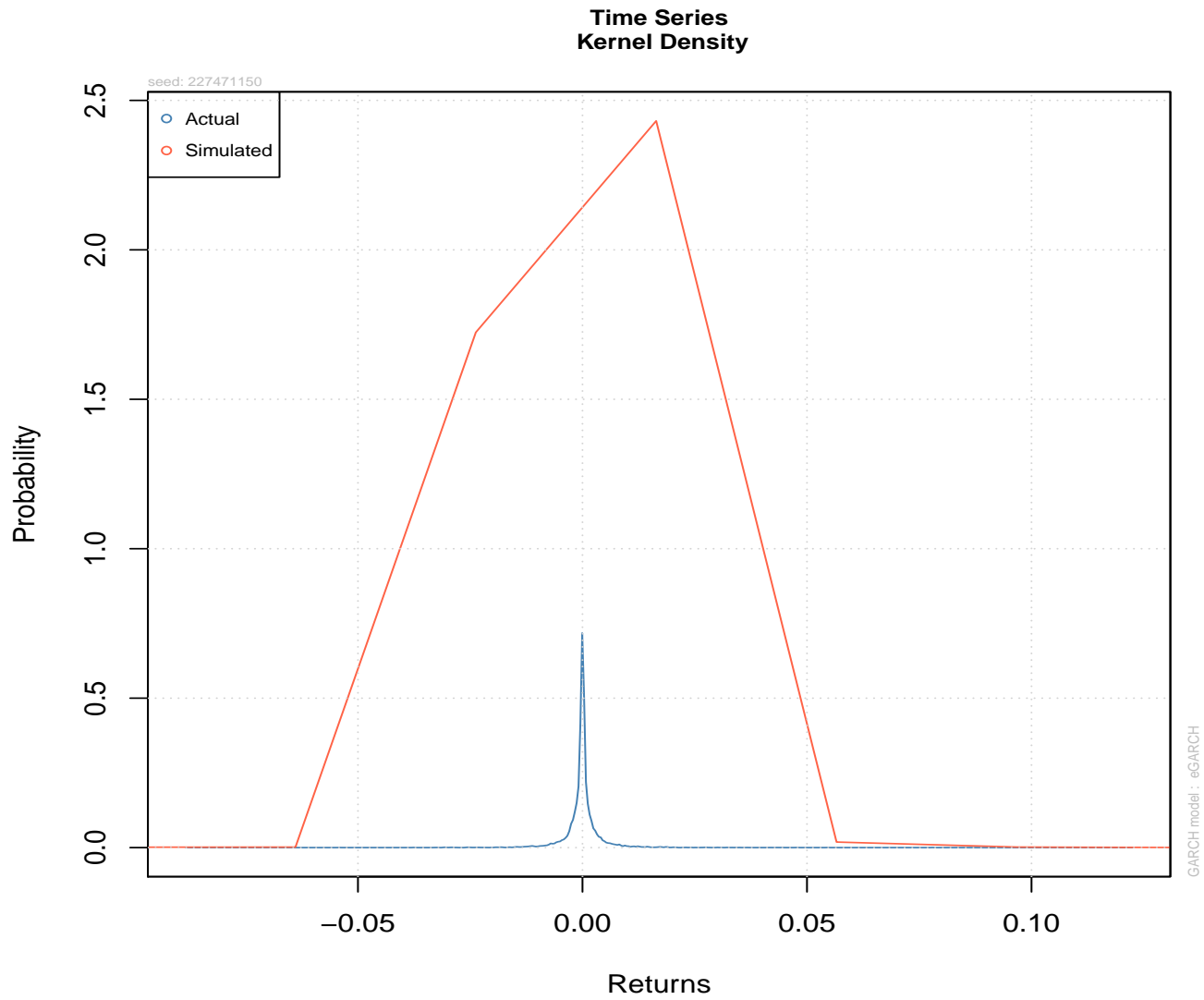Figure 4.19: ARMA(2,0)/EGARCH (1,1) Conditional SD Simulation Density
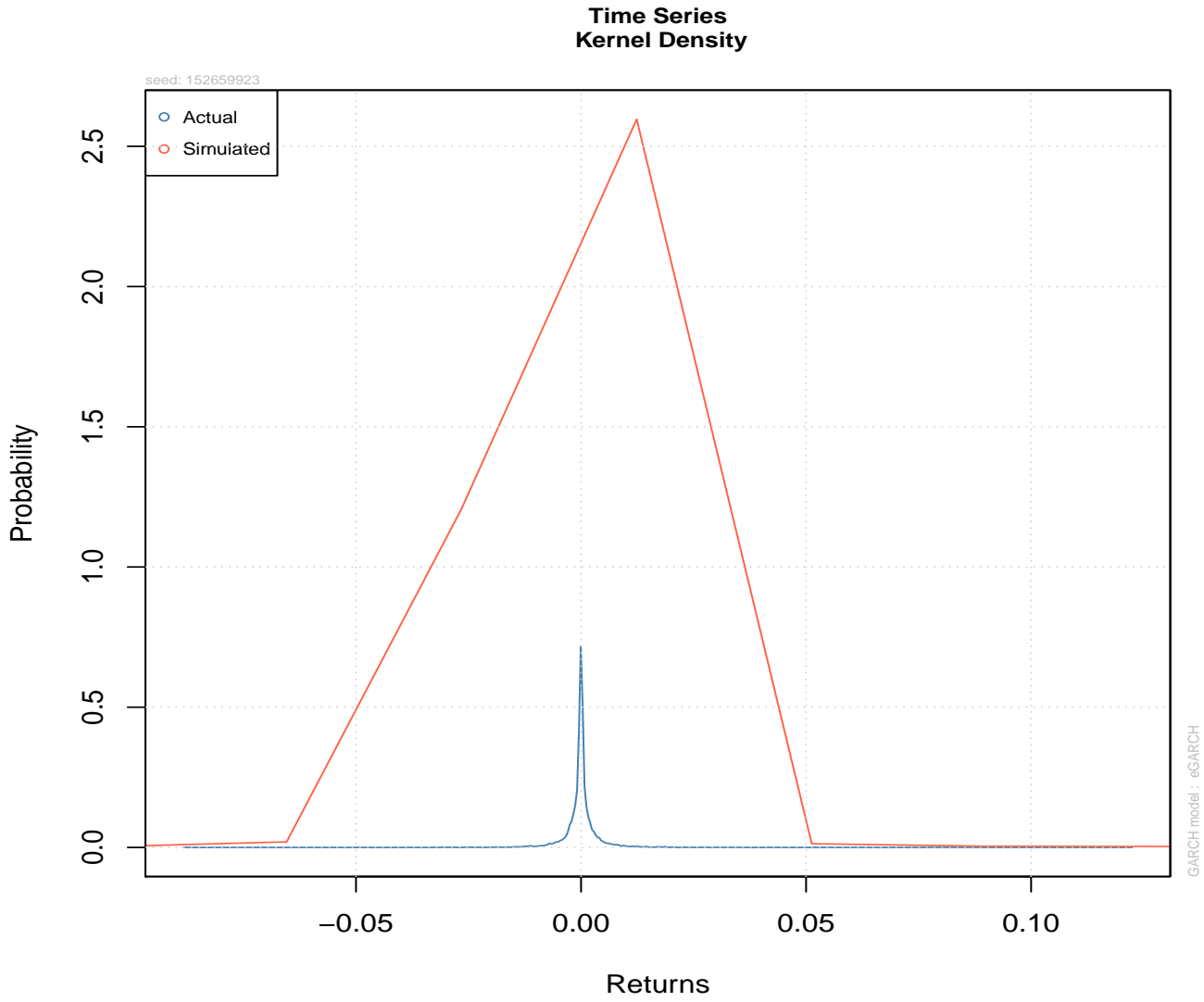
Figure 4.20: ARMA(2,0)/EGARCH (1,1) Return Series Simulation Path Density
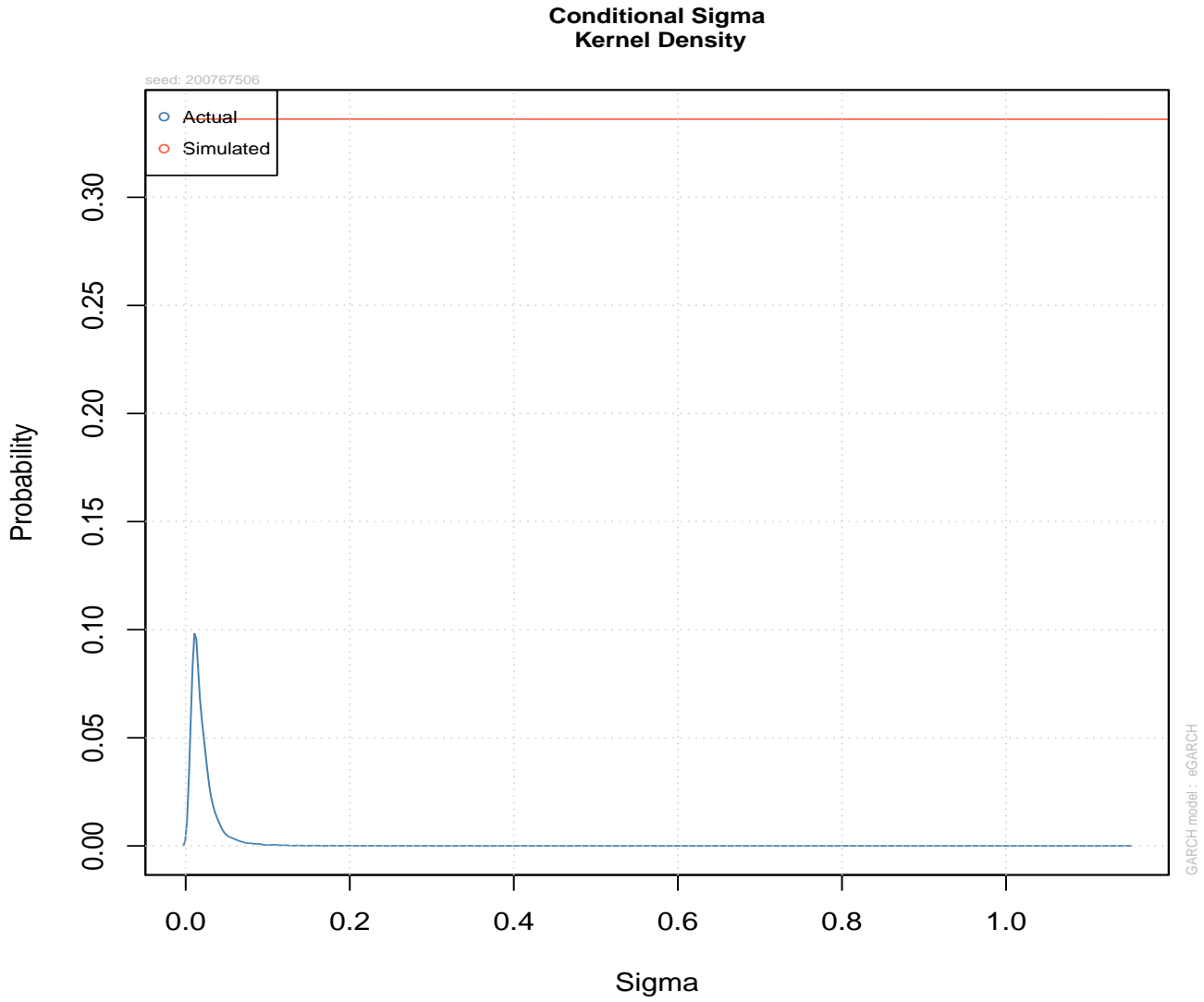
Figure 4.21: ARMA(0,2)/EGARCH (1,1) Conditional SD Simulation Density

Figure 4.22: ARMA(0,2)/EGARCH (1,1) Return Series Simulation Path Density

Figure 4.23: ARMA(1,1)/EGARCH (1,1) Conditional SD Simulation Density

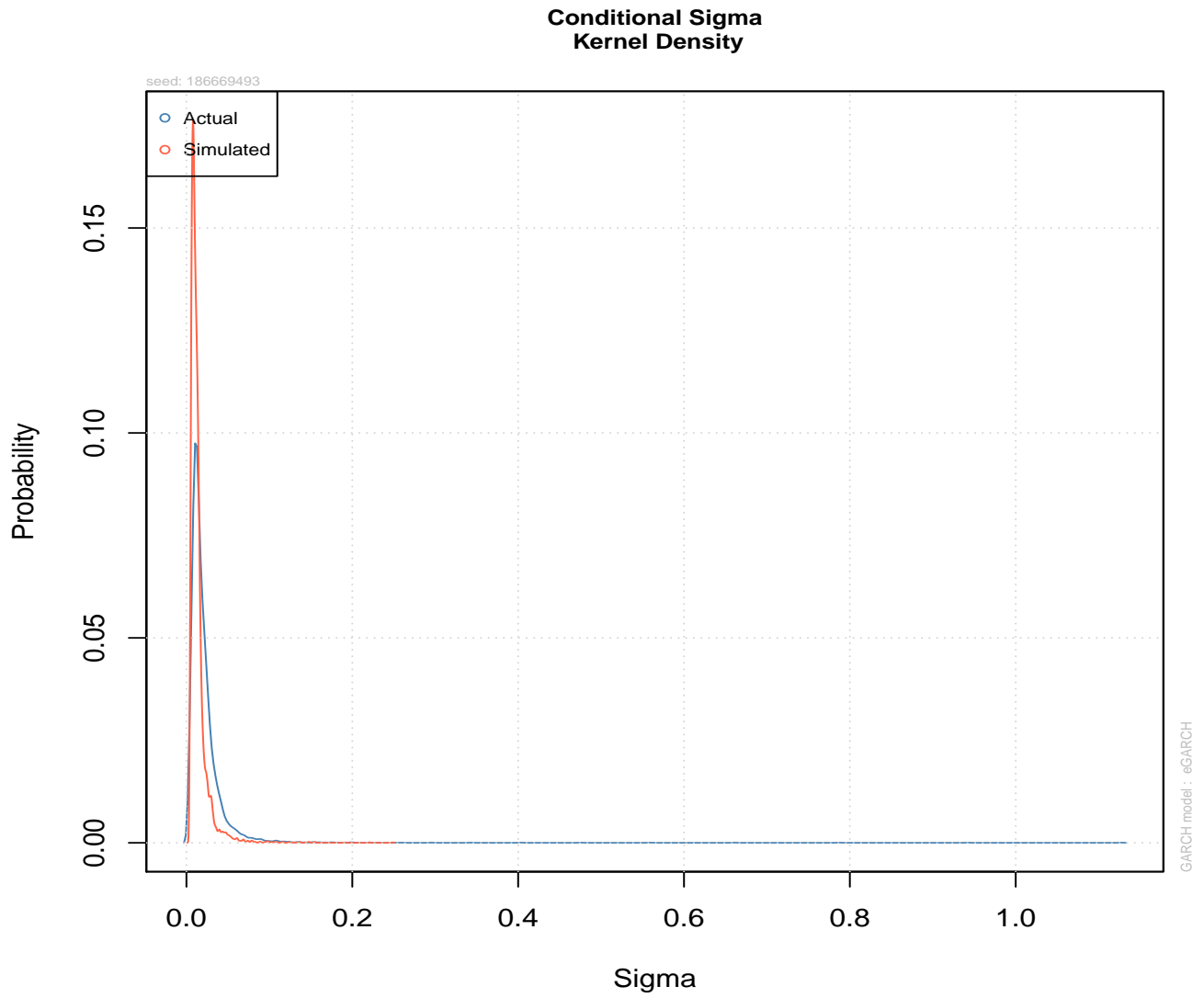Figure 4.24: ARMA(1,1)/EGARCH (1,1) Return Series Simulation Path Density

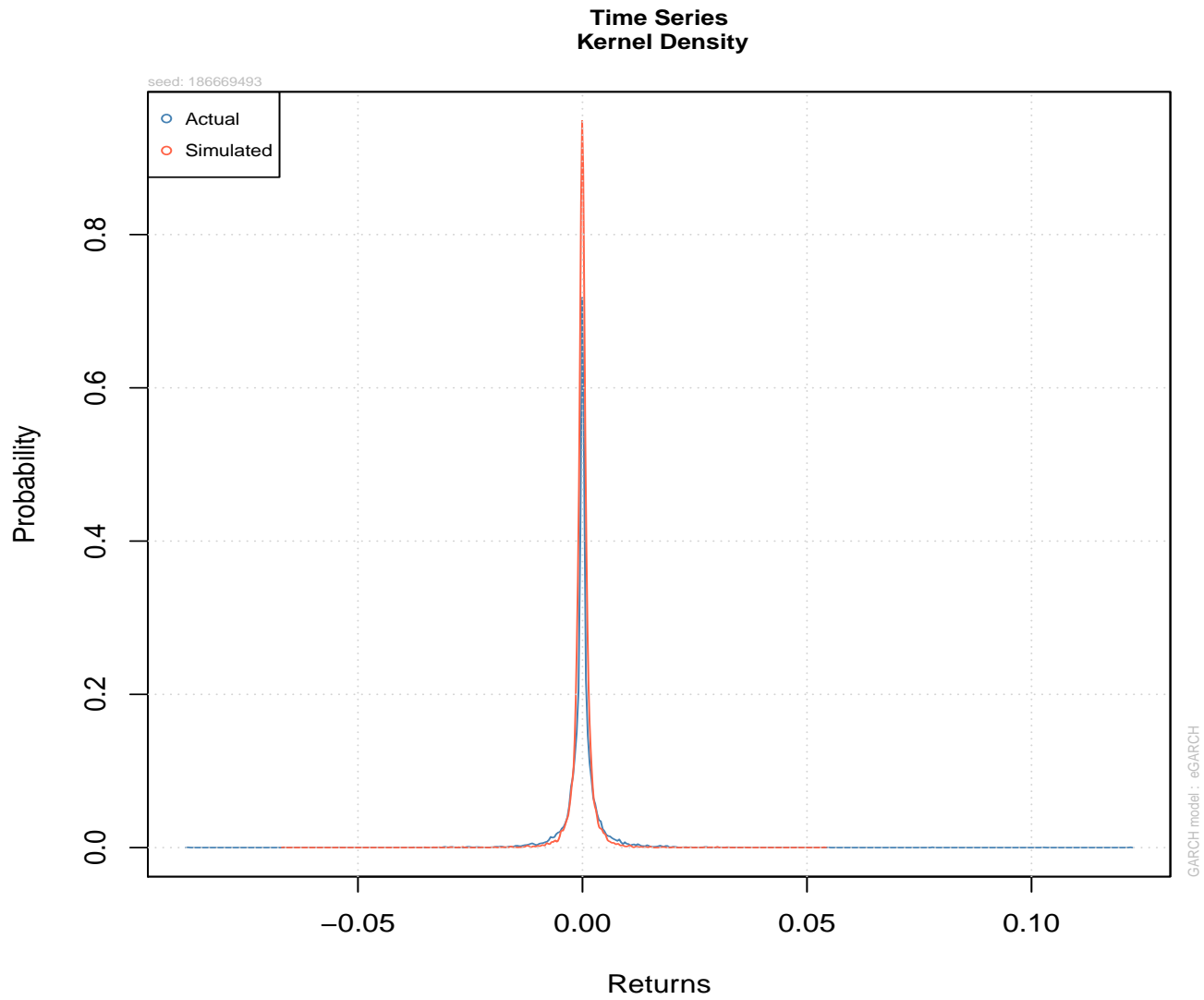Figure 4.25: ARMA(1,0)/EGARCH (1,1) Conditional SD Simulation Density

Time Series
Kernel Density

seed: 186669493

Actual
Simulated

GARCH model : eGARCH

Probability

Returns

Figure 4.26: ARMA(1,0)/EGARCH (1,1) Return Series Simulation Path Density
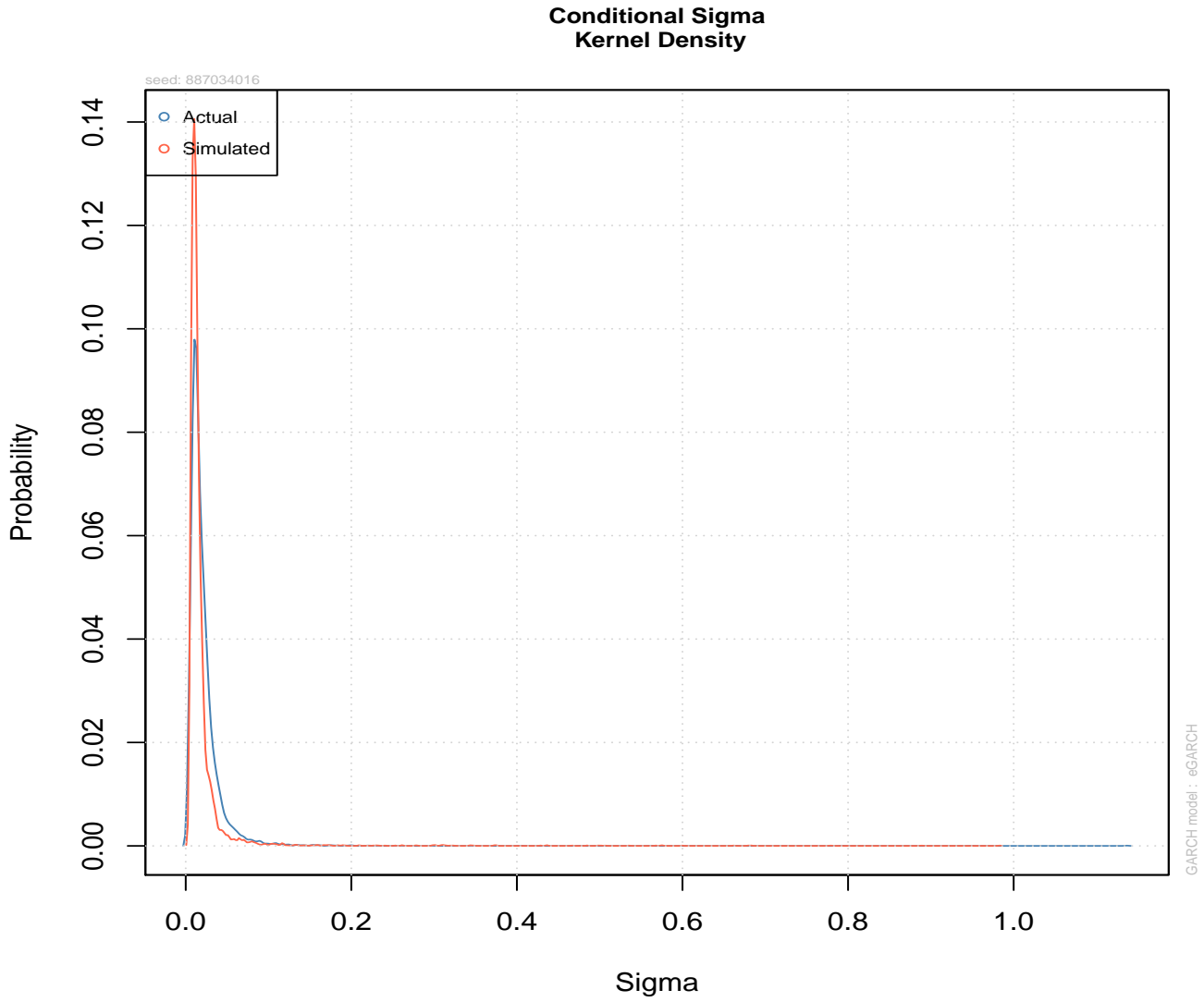
63

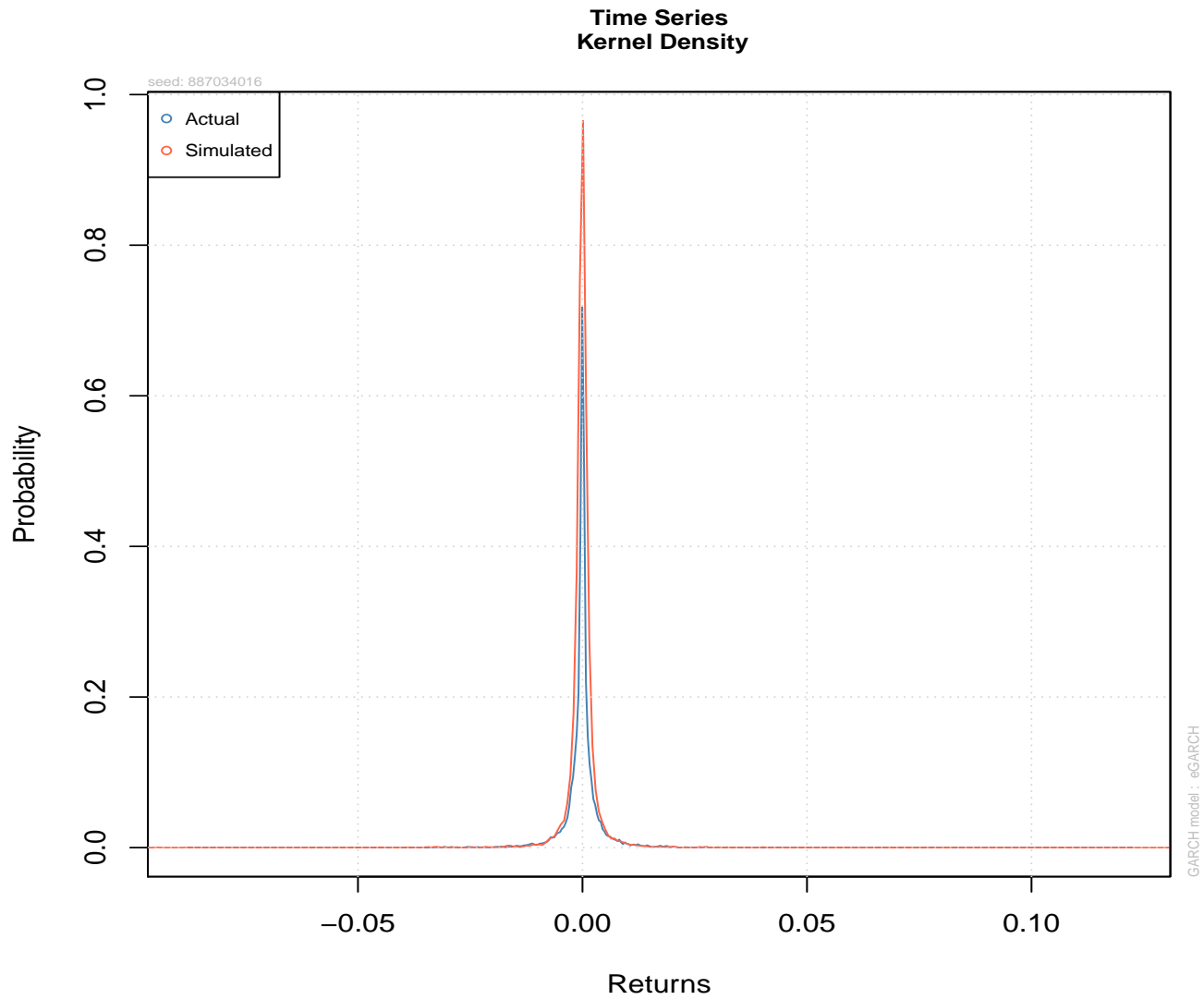Figure 4.27: ARMA(0,1)/EGARCH (1,1) Conditional SD Simulation Density

Figure 4.28: ARMA(0,1)/EGARCH (1,1) Return Series Simulation Path Density
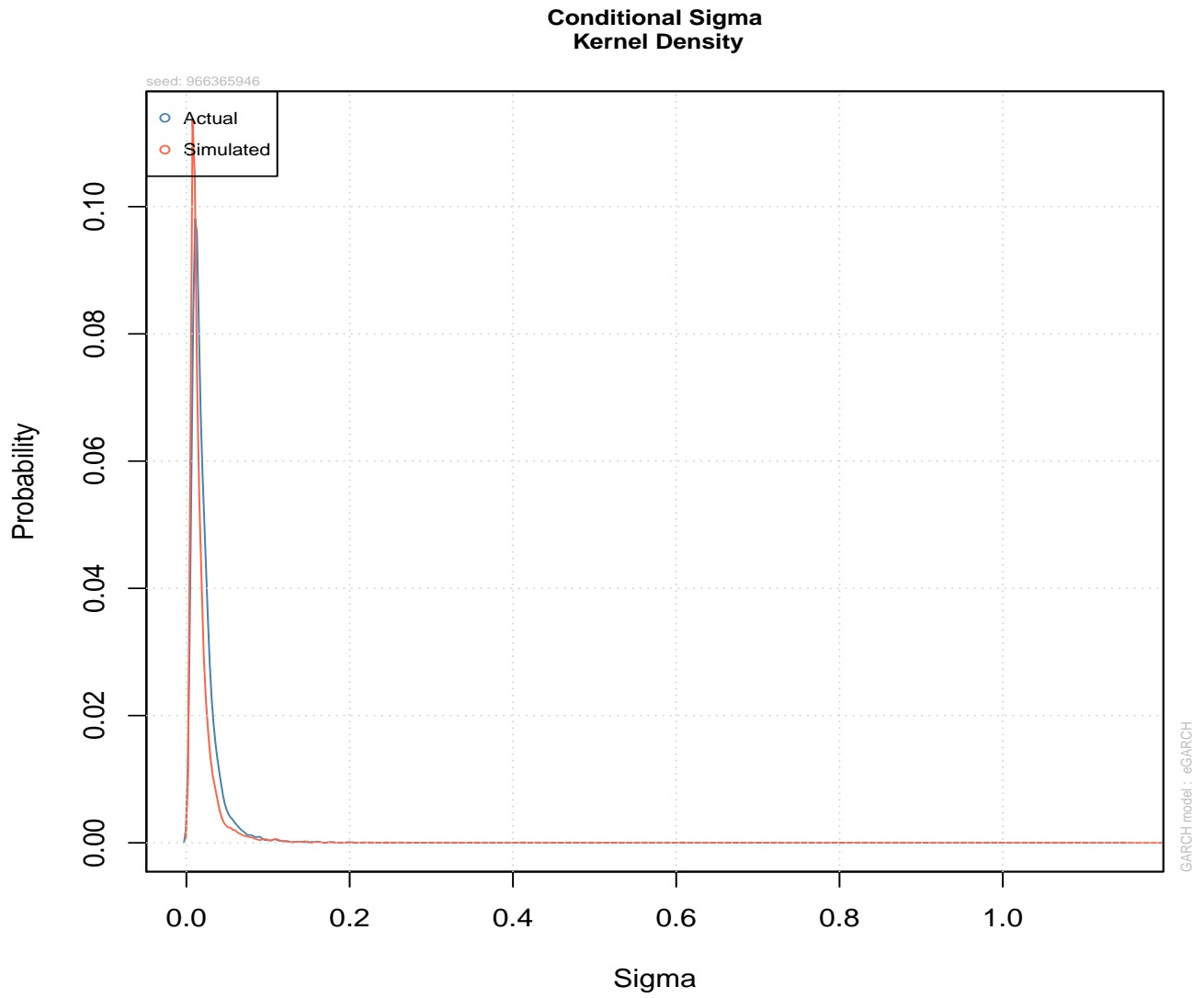
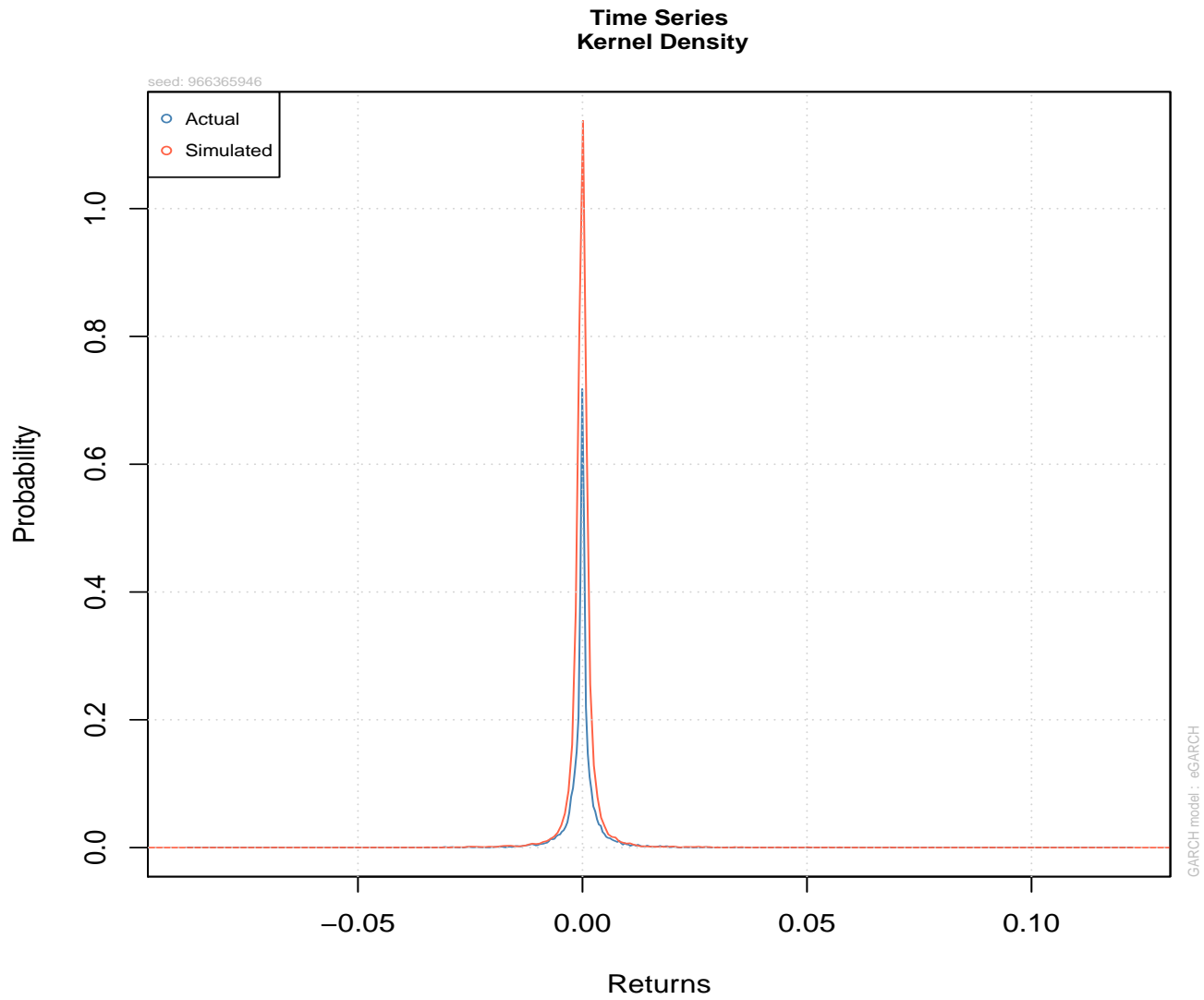Figure 4.29: ARMA(1,2)/EGARCH (1,1) Conditional SD Simulation Density

Figure 4.30: ARMA(1,2)/EGARCH (1,1) Return Series Simulation Path Density

From the above plots of simulated vs actual data, ARMA(0,1)/EGARCH(1,1), ARMA(1,0)/EGARCH(1,1) and ARMA(1,2)/EGARCH(1,1) appear to provide equally good fits with less parameters. We opt for the ARMA(1,0)/EGARCH(1,1) over the previous ARMA(2,2)/EGARCH(1,1) in line with the principle of parsimony and because the fit seems good enough.

4.2.7 Model Analysis

Conditional Variance Dynamics
GARCH Model : eGARCH(1,1)
Mean Model : ARFIMA(1,0,0)
Distribution : sstd

Table 4.7: Optimal Parameters

	Estimate	Std. Error	t value	Pr($> |t|$)
mu	-0.000044	0.000003	-13.9209	0.00000
ar1	0.089802	0.009078	9.89279	0.00000
omega	-0.194204	0.003058	-63.18771	0.00000
alpha1	0.013307	0.059055	0.22533	0.82172
beta1	0.977204	0.000039	25250.67302	0.00000
gamma1	1.701140	0.067519	25.19494	0.00000
skew	0.994035	0.006665	149.13891	0.00000
shape	2.010000	0.000415	4841.74253	0.00000

gamma1 is positive and statistically significant at 1% level of significance. The magnitude is 1.701140 and the sign is positive. The positive value shows that past positive events have more influence on future volatility. The positive and statistically significant coefficient shows that the KES/USD exchange rate volatility is higher after positive shock i.e. bad news generates less volatility than good news. In financial markets, depreciation of currency is usually followed by higher volatility, but this is not the case for the KES/USD exchange rates.

The sum of $\alpha + \beta = 0.990511$, which is less than 1. This shows that volatility is persistent.

The overall model can now be written as:
Mean Equation: $X_t = Z_t + 0.08902Z_{t-1}$

Variance Equation: $ln(\sigma_t^2) = -0.194204 + 0.977204 ln(\sigma_{t-1}^2) + 0.013307 \left[|\frac{\epsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}}| - E(\epsilon_{t-1}) \right] + 1.701140 \frac{\epsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}}$

4.2.8 Residual Analysis

Table 4.8: Q-Statistics on Standardized Residuals

	statistic	p-value
Lag[1]	32.38	1.266e-08
Lag[p+q+1][5]	33.51	7.072e-09
Lag[p+q+5][9]	59.68	1.417e-11

H_0 : No serial correlation

Table 4.9: Q-Statistics on Standardized Squared Residuals

	statistic	p-value
Lag[1]	0.0004202	0.9836
Lag[p+q+1][3]	0.0512348	0.8209
Lag[p+q+5][7]	0.2522366	0.9984
degrees of freedom = 2		

From Table 4.8, the high p-values lead us to accept the null hypothesis of No serial correlation. This further strengthens the observation of the ACF of squared standardized residuals in Figure 4.31 below:
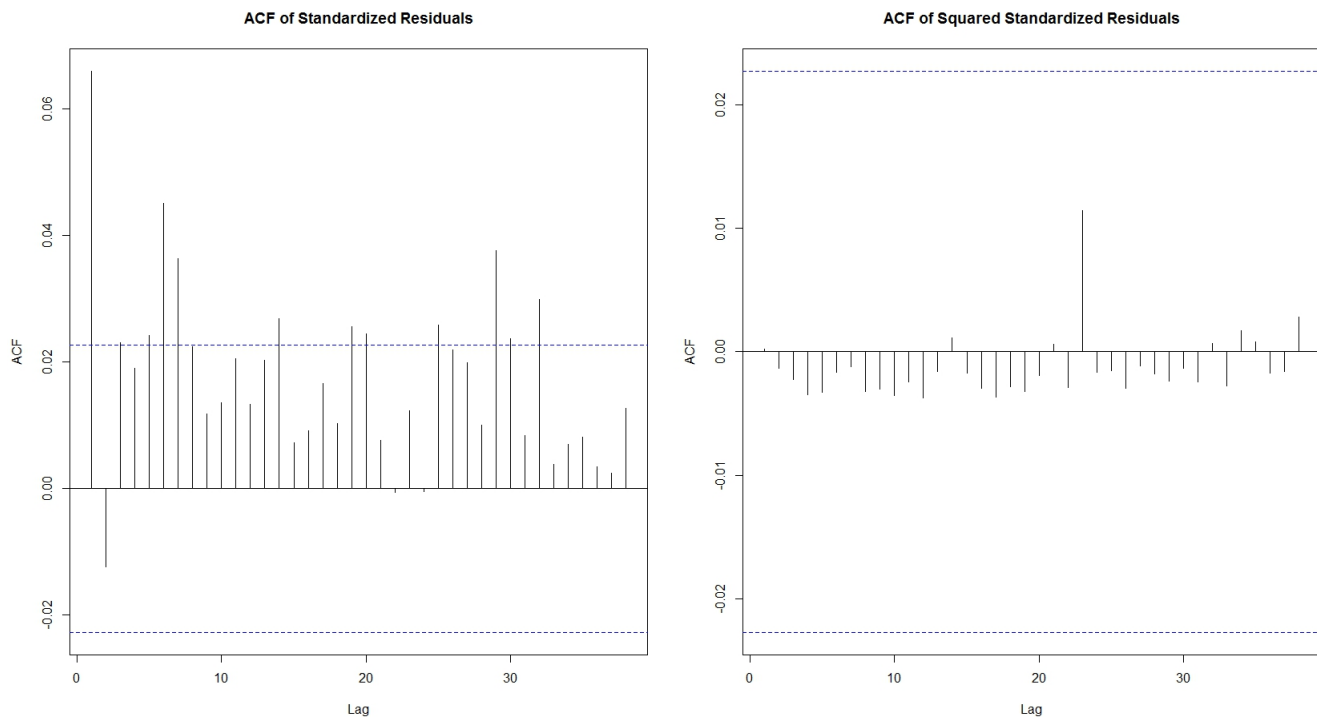

Figure 4.31: ACF of Standardized Residuals and Squared Standardized Residuals

Table 4.10: ARCH LM Tests

	Statistic	DoF	P-Value
ARCH Lag[2]	0.0136	2	0.9932
ARCH Lag[5]	0.2204	5	0.9989
ARCH Lag[10]	0.4980	10	1.0000

The ARCH LM tests the null hypothesis that there are no more ARCH effects in the residuals. From the p-values as listed in Table 4.9, at 1%,5% and even 10% level of significance, we can conclude that there are no more ARCH effects in the residuals which indicates that the volatility model is correctly specified.

4.3 VOLATILITY ANALYSIS OF FOUR PERIODS

The data is further split into four periods i.e. 1993-1999, 2000-2004, 2005-2009, 2010-2014. The plots of the original series vs their volatility are shown below:
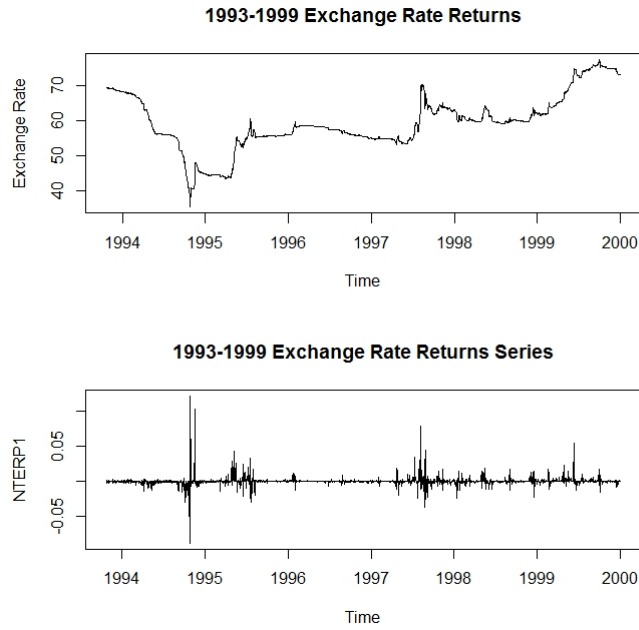

Figure 4.32: 1993-1999

In the period 1993-1999,we can see from Figure 4.32 that towards the end of 1994 and 1997, there were higher spikes in the exchange rate returns series as compared to the rest of the years.
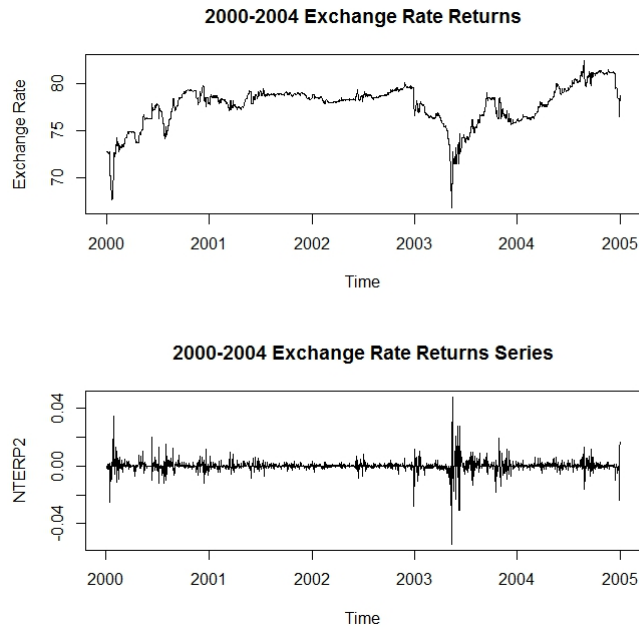

Figure 4.33: 2000-2004

In the period 2000-2004, again increase in the fluctuations as indicated by the high spikes are seen at the beginning of the year 2000 and mid 2003.
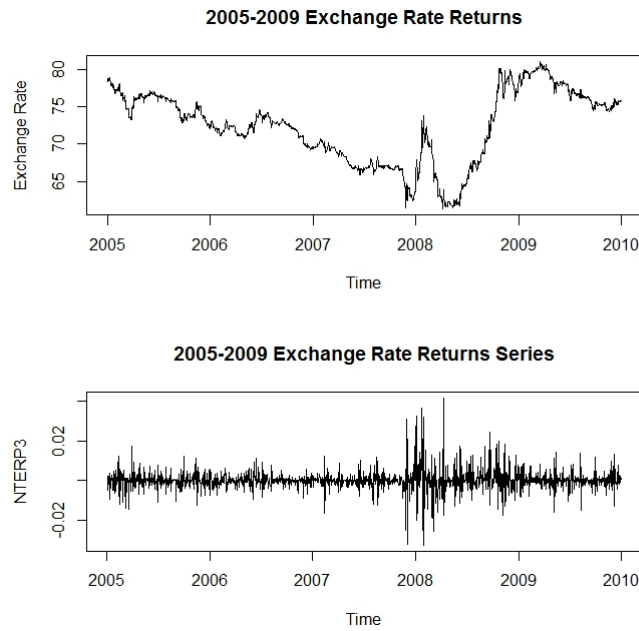

Figure 4.34: 2005-2009

Between the years 2005 to 2009, again high and erratic spikes are seen towards the end of 2007 all the way to the beginning of 2009.
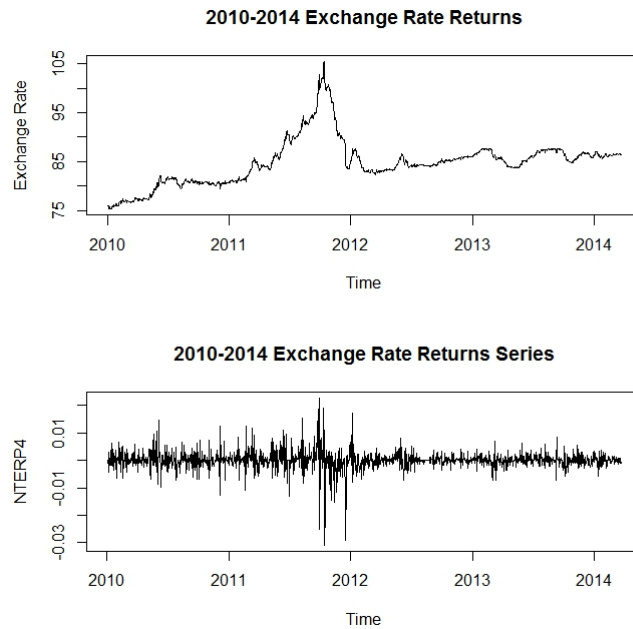

Figure 4.35: 2010-2014

The last period 2010 to 2014 has significant spikes towards the end of 2011 and terminate as 2012 begins.

The spikes and dips represent volatility of the KES/USD exchange rate return series. An important observation made is that the spikes in volatility are attributed to an appreciation of the KES against the USD. 1995 saw the dollar at the lowest rate ever attained. Mid 1997 when compared to the exchange rate returns also shows an appreciation of the currency.

The years 2000 and mid 2003 also show high spikes compared to the rest of the period. These also correspond to an appreciation of the KES. The year 2000 marked the crossing over into a new millenium and many currencies were affected negatively yet for the KES, the beginning of that year saw a steady appreciation of the currency.

The period 2005-2009 had erratic spikes clustered towards the end of 2007 upto a few months into 2008 where it reached it's highest at 73,followed by a persistence in the volatility. Thereafter, it started to appreciate once more to a low of around 63 where we see the single outstanding volatility spike. The significant occurrence in that period in Kenya was elections in December of 2007, which then culminated in post election violence which spilled over to the beginning of 2008. It can be observed from the exchange rate returns that the exchange rate had well appreciated prior to the elections and hence the onset of the volatility spikes, this persists for a while as the exchange rate begins to appreciate but then dies down. CBK attributed the depreciation of the KES/USD exchange rate to increased demand for the dollar driven by expectations of increased importation of maize.(Kenya Monthly Economic Review, December 2008)

In the period 2010-2014 another significant occurrence in the Kenyan economy was the greatest depreciation the currency has ever faced. The KES/USD exchange rates went as high as 105 in October which saw the Central bank of Kenya increase lending rates in a bid to curb the runaway inflation. The difference with this volatility spike is that it occurred after a depreciation of the KES, contrary to the spikes in the other periods which occurred after a currency appreciation. This could easily be explained by the fact that in a span of 148 days i.e. from 07/06/2011 to 27/10/2011, the exchange rate ranged from 86-105 then back to 100, which brought about the high fluctuation.

Towards the end of 2011, there is another outstanding spike which is associated with a currency appreciation. When the KES/USD exchange rate hit the all time high of 105, no significant volatility spikes are observed but when the exchange rate starts coming down, a significant spike is observed just at the beginning of 2012. The year 2011 remarkable event in the Kenyan economy was the shilling crash. The CBK attributed the exchange rate appreciation to "tightening of Monetary Policy stance in November 2011 through June 2012 to reduce inflationary expectations and exchange rate volatility".(Kenya Monthly Economic Review, December 2012).

In general, we can conclude that the volatility spikes in the KES/USD exchange rates occur prior to and just before appreciation of the currency.

Chapter 5

CONCLUSION

Information on volatility of exchange rates is crucial to various groups such as importers, exporters, investors, policy makers, Governments etc. There is thus need to build models that can then be used for simulation and possibly forecasting. Stylized facts about financial time series such as volatility clustering and persistence were observed. Normality assumption was rejected in the original exchange rates data as well as in the residuals of the fitted model. This was inferred from the QQ-plots, histogram, JB statistic as well as the skewness and kurtosis coefficients. The chosen model with the skewed student-t distribution was better suited to accommodate the skewness and kurtosis in the exchange rates return series.

This study opted for lower specifications of the GARCH models in spite of existence of higher orders because empirical evidence shows that the lower specifications are able to sufficiently capture the characteristics of exchange rates while at the same time upholding the principle of parsimony (Kocenda and Valachy, 2006)

Various ARMA(p_A, q_A)/GARCH(1,1) and ARMA(p_A, q_A)/EGARCH(1,1)models were fitted with variations being made to the parameters of the mean ARMA model as well as in the conditional distributions used i.e. normal, skewed normal, student-t and skewed student-t. The volatility model on the other hand was retained at (1,1) since many studies have shown that there is not much difference between higher order models so the GARCH(1,1) model is sufficient and obeys the principle of parsimony. The best fitting model was selected based on the lowest AIC and the highest log-likelihood. The resultant model was found to be an ARMA(2,2)/EGARCH(1,1) with a skewed student-t conditional distribution.

The test for existence of asymmetry in the KES/USD was compelled by recent empirical evidence of strong support for its existence in foreign exchange markets (Firdmuc & Horvath, 2007). The EGARCH model provides a better fit than the GARCH model and its advantages over the GARCH model are that first, it can capture leverage effects and secondly, that there is no restriction that the parameters α_1 and β_1 must be positive. (Hansen and Lunde, 2005; Andersen, Bollerslev, Chou and Kroner,1992) are of a different opinion;that foreign exchange returns usually exhibit symmetric volatility unlike equity markets. They say that past positive and negative shocks have the same effects on future volatility. Bollerslev, Chou and Kroner (1992) argue that, "whereas stock returns have been found to exhibit some degree of asymmetry, the two sided nature of foreign exchange markets makes such asymmetries less likely."

The fact that $\gamma \neq 0$ leads us to conclude that the exchange rates return series exhibits some leverage effect. For the KES/USD, this value is positive meaning that a positive shock has more impact on exchange rate volatility than a negative shock. This is in contrast to the leverage effects results in the developed countries. The EGARCH model was able to capture this.

Using AIC and LLH, the ARMA(2,2)/EGARCH(1,1) model with student-t conditional distribution for the errors was selected. For the sake pf parsimony, other lower lag models were considered to see if any of them that was rejected by AIC could still give a good fit. Kernel density plots of the conditional standard deviation simulation and return series simulation path densities were observed and the one with the closest fit of actual vs simulated observation was selected. This turned out to be the ARMA(1,0)/EGARCH(1,1) model. i. e. AR(1)/EGARCH(1,1). The EGARCH model was also found to be most suitable in the works of Miron et al.(2010), Zahangir et al.(2012), Ebru et al.(2013), Kamal et. al (2011) among others. In the Kenyan market, Rotich(2014) also found that the AR(1)/EGARCH(1,1) with student-t distribution was the best model for the KES/USD exchange rates.

The data was further split into four sub periods and volatility in each period is examined separately. The periods are as follows:$1993 - 1999$ with 2261 observations, 2000-2004 with 1827 observations, 2005-2009 with 1826 observations and 2010-2014 with 1541 observations. Each of the four periods has an election year included. Spikes in volatility are observed at times when there is an appreciation of the KES/USD currency further confirming leverage effects.

5.1 Recommendation

1. Use of Bayesian statistics in modelling exchange rate volatility.

2. A study with other asymmetric GARCH models.

Bibliography

[1] Abdalla, S. Z. S. (March 2012). Modelling exchange rate volatility using garch models: Empirical evidence from arab countries. *International Journal of Economics and Finance Vol. 4, No. 3*.

[2] Adedayo A. Adepoju, O. S. Y. and O. D. Ojo (2013). Estimation of garch models for nigeria exchnage rates under non-gaussian innovations. *Journal of Economics and sustainable development Vol. 4,No.3*.

[3] Alam, M. Z. and M. A. Rahman (2012). Modelling volatility of the bdt/usd exchange rate with garch model. *International Journal of Economics and Finance Vol. 4, No. 11*.

[4] Alexander, C. and E. Lazar (2006). Normal mixture garch (1, 1): Applications to exchange rate modelling. *Journal of Applied Econometrics Vol 21*, pp. 307–336.

[5] Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Economics 31*, 307–327.

[6] CBK. Kenya monthly economic review. *CBK Website*.

[7] Chaiwat Kosapattarapim, Y.-X. L. and M. McCrae. Evaluating the volatility forecasting performance of best fitting garch models in emerging asian stock markets.

[8] Chiarah, B. I. A. (September 2013). Sources of real exchange rate volatility in the ghanaian economy. *Journal of Economics and International Finance Vol. 5(6)*, PP. 232–238.

[9] Chipili, J. M. Modelling exchange rate volatility in zambia. *Department of Economics, University of Leicester, Leicester,LEI 7RH, U.K*.

[10] D Ng Cheong Vee, P. N. G. and N. Sookia (2011). Forecasting volatility of usd/mur exchange rate using a garch (1, 1) model with ged and student's-t errors. *University of Mauritius Research Journal 17*.

[11] Danson Musyoki, G. P. P. and M. Pundo (2012). Real exchange rate volatility in kenya. *Journal of Emerging trends in Economics and Management Sciences (JETEMS) 3 (2)*, pp. 117–122.

[12] David McMillan, A. S. and O. Gwilym (2000). Forecasting uk stock market volatility. *Applied Financial Economics 10*, 435–448.

[13] Dima Alberg, H. S. and R. Yosef (2008). Estimating stock market volatility using asymmetric garch models. *Applied Financial Economics 18*, 1201–1208.

[14] Dominguez, K. M. (1998). Central bank intervention and exchange rate volatility. *Journal of International Money and Finance*, PP. 161–190.

[15] Ebru Caglayan, T. U. and T. Dayioglu (Septemberr 2013). Modelling exchange rate volatility in mist countries. *International Journal of Business and Social Science Vol. 4 No. 12*.

[16] Ederington, L. H. and W. Guan (September 2004). Forecasting volatility.

[17] Emenike, K. O. (2010). Modelling stock return volatility in nigeria using garch models. *Munic Personal RePEc Archive Paper No. 22723.*

[18] Engle, R. E. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation. *Econometrica Vol. 50, N0. 4*, pp. 987–1007.

[19] Eriksson, K. (2013). On return innovation distribution in garch volatility modelling.

[20] Firdmuc, J. and R. Horvath (October 2007). Volatility of exchange rates in selected new eu members: Evidence from daily data. *CESIFO working paper No. 2107 Category 6: Monetary Policy and Internatiional Finance.*

[21] Granger, S.-H. P.-H. P. J. (June 2003). Forecasting volatility in financial markets. *Journal of Economic Literature Vol. XLI*, pp. 478–539.

[22] Insah, B. (2013). Modelling exchange rate volatility in a developing country. *Journal of Economic and sustainable development Vol. 4, No. 6.*

[23] Isaya Maana, P. N. m. and R. Odhiambo (July 2010). Modelling the volatility of exchange rates in the kenyan market. *African Journal of Business Management Vol. 4(7)*, pp. 1401–1408.

[24] John Dukich, K. Y. K. and H.-H. Lin (December 2010). Modelling exchange rates using the garch model.

[25] K, S. and D. G. Shanmugasundaram. Foreign exchange rate volatility of indian rupee/us dollar.

[26] Kipkoech, R. T. (2014). Modeling volatility under normal and student-t distributional assumptions (a case study of the kenyan exchange rates). *American Journal of Applied Mathematics and Statistics Vol. 2, No. 4, pp 179-184.*

[27] Latifa S. Omar, J. K. K. and J. M. Mutiso (2013). Garch modelling in monthly foreign exchange in kenya. *Americal journal of Mathematical Science and Applications Volume 1, Number 1*, PP. 13–39.

[28] Miron, D. and C. Tudor (2010). Asymmetric conditional volatility models: Empirical estimation and comparison of forecasting accuracy. *Romanian Journal of Economic Forecasting 3.*

[29] Neely, C. J. and P. A. Weller. Predicting exchange rate volatility: Genetic programming versus garch and riskmetrics.

[30] O., E. K. (May 2010). Modelling stock returns volatility in nigeria using garch models. *MPRA Munic Personal RePEc Archive Paper No. 22723.*

[31] Omolo, S. A. (2012). The link between exchange rates and interest rates. *Post Graduate Diploma Project.*

[32] Poon, S.-H. (August 11, 2008). Arch.

[33] Posedel, P. (2005). Properties and estimation of garch (1, 1) model. *Metodoloski Zvezki Vol. 2, No. 2*, pp. 243–257.

[34] Yasir Kamal, Hammad-Ul-Haq, U. G. and M. M. Khan (February 2012). Modelling exchange rate volatility using generalized autoregressive conditional heteroscedastic (garch) type models. *African Journal of Business Management Vol.6(8)*, pp. 2830–2838.

[35] Yu, J. (2002). Forecasting volatility in the new zealand stock market. *Applied Financial Economics 12*, pp. 193–202.