UNIVERSITY OF NAIROBI

SCHOOL OF MATHEMATICS

Application of GARCH to Model the KES/US$ Foreign Exchange Rates Returns

This research project is submitted to the School of Mathematics of the University of Nairobi in partial fulfillment of the requirement for the degree of Masters of Science in Social Statistics.

By

DOUGLAS O. BARASA

DECLARATION

This project as presented in this report is my original work and has not been presented for any other university award

Student: Douglas Otunga Barasa Reg. No.: I56/78047/09

Signature: ________________ Date: 01/08/2011

This project has been submitted as a partial fulfillment of the requirements for Masters of Science in Social Statistics of the University of Nairobi with our approval as the university supervisors:

Mr. John M. Ndiritu Signature Date: 01/08/2011

Dr. Joseph I. Mwaniki Signature Date: 01/08/2011

ACKNOWLEDGEMENT

I acknowledge the instrumental assistance accorded by my supervisors Mr. John M. Ndiritu and Dr. Joseph I. Mwaniki both of University of Nairobi's School of Mathematics and the Central Bank of Kenya for availing the data. I also wish to acknowledge the R development core team 2011 for the R statistical software used for this research project for free and developers of the various packages.

ABSTRACT

The USA dollar (US$) is the most prominent currency around the world for transactions and also as foreign reserves for many Central Banks, for example in 2006, Central Bank of Kenya had 52% of its foreign reserves in US dollar currency. The change in the strength of the US$ relative to Kenya shilling (KES) has an impact on many socio-economic sectors in Kenya. This research focuses on KES against US$ daily exchange rate returns from 2nd November 2004 to 31st December 2010. The Box-Jenkins models for time series assume homoscedasticity in the time series, however the returns exhibits *stylized facts* which can only be well modeled using conditional heteroscedacity-type of models. This project considers the application of Autoregressive Integrated Moving Average (ARIMA) models on the exchange rate. The ARIMA (4,1,2) model was fitted and its residuals exhibited volatility clustering and hence Generalized Auto regression Conditional Heteroscedacity (GARCH) was applied to address these characteristics. A quasi maximum likelihood estimation procedure was used and the estimators given. It was found that the returns are leptokurtic and have fat tails. GARCH(1,1) were fitted on the returns and was found to fit the returns well and its residuals found to be white noise and homoscedastic. The one day ahead forecasting are quite good implying that it could be used for future prediction on the volatilities of the returns.

Table of Contents

List of figures

List of tables

List of Pictures

ABBREVIATIONS AND ACRONYMS

ARCH	Auto Regression Conditional Heteroscedacity
BOVESPA	Brazil Sao Paulo Stock Exchange Index
CBK	Central Bank of Kenya
Forex	Foreign Exchange
GARCH	Generalized Auto Regression Conditional Heteroscedacity
i.i.d.	Independent and identically distributed
KES/KSh	Kenya Shilling
US$	United States Dollar
MIBOR	Mumbai Inter-Bank Offer Rate
NARC	National Alliance Rainbow Coalition
S&P 500	Standards and Poor 500 Index
$Y \sim N(0,1)$	Y is normally distributed with mean 0 and unit variance

CHAPTER 1

INTRODUCTION

The sequence of random Variables $\{ Y_t : t = 0, \pm 1, \pm 2, \pm 3 \ldots \}$ is called a stochastic process and serves as a model for an observed time series. In other words, time series is an ordered sequence of values (data points) of variables at equally spaced time intervals. It is known that the complete probabilistic structure of such a process is determined by the set of distributions of all finite collections of the Y's.

Time series is highly applied in economics, marketing, demography among others. A time series is made up of several components which include; trend, seasonal variations, cyclic variation and random movements.

The main objectives of time-series analysis are:

- Description. To describe the data using summary statistics and/or graphical methods. A time plot of the data is particularly valuable.
- Modelling. To find a suitable statistical model to describe the data generating process. A univariate model for a given variable is based only on past values of that Variable, while a multivariate model for a given Variable may be based, not only on past values of that Variable, but also on present and past values of other (predictor) Variables. In the latter case, the Variation in one series may help to explain the Variation in another series. Of course, all models are approximations and model building is an art as much as a science.
- Forecasting. To estimate the future values of the series. Most authors use the terms 'forecasting' and 'prediction' interchangeably and we follow

this convention. There is a clear distinction between steady-state forecasting, where we expect the future to be much like the past, and What-if forecasting where a multivariate model is used to explore the effect of changing the policy Variables.

- Control. Good forecasts enable the analyst to take action so as to control a given process, whether it is an industrial process, or an economy or whatever. This is linked to "What if…" forecasting.

1.1 International role of the US Dollar

The dollar is the most prominent currency in the world. It plays a central role in international trade and finance as both a store of value and a medium of exchange. Many countries maintain an exchange rate regime that anchors the value of their home currency to that of the dollar. Dollar holdings make up a large share of official foreign exchange reserves, the foreign currency deposits and bonds maintained by central banks and monetary authorities. In international trade, the dollar is widely used for invoicing and settling import and export transactions around the world[1]. In 2009, the dollar assets accounted for about two-thirds of the reserve assets of industrialized and developing countries.

In 2006 the Central Bank of Kenya (CBK) held 52% of its foreign currency reserves in US Dollar currency. CBK licenses and monitors activities of all the agencies that do forex trading which include banks, hotels and forex bureaus that are found in major towns in the country. The exchange rate is important because it allows the conversion of Kenya shilling into foreign currencies thus facilitating purchase of goods and services from other countries (CBK, 2008).

Most businesses in Kenya such as manufacturing, stock exchange, tourism, petroleum who import or export products and services depend a lot on the KES/US$ exchange rates as it is the accepted global exchange currency. A weak shilling makes Kenyan goods and services cheaper in the international market but makes

[1] www.voxeu.org/index.php?q=node/4819

imports more expensive. So exporters benefit while importers lose (i.e. they import less goods from abroad). Conversely, a strong shilling makes our goods and services expensive in the international market and makes our imports more affordable. In this case, importers gain while the exporters lose (CBK, 2008). It's hence imperative to have an excellent time series model to approximate the exchange rate value at a time t in the future putting into consideration the prevailing world conditions. With this significant share it means that the country's business transactions are pegged on the dollar, as any depreciation or appreciation has an impact on the Kenya shilling.

Holmes (2003) found that exchange rate depreciation was inflationary although the impact could not prevail over the gains from increased external competitiveness. Depreciation reduces the real value of assets denominated in the local currency and increases the real value of foreign currency denominated assets. Assuming a constant money supply, domestic inflation increases if the first effect dominates the second effect.

When price change is defined relative to some initial price, it is known as a return. In this research we will measure change in value of a portfolio (often referred to as the adverse price move) in terms of log price changes also known as continuously compounded returns.

1.2 Determinants of Exchange Rates

Numerous factors determine exchange rates, and all are related to the trading relationship between two countries. Exchange rates are relative, and are expressed as a comparison of the currencies of two countries.

The following are some of the principal determinants of the exchange rate between two countries. Note that these factors are in no particular order; like many aspects of economics, the relative importance of these factors is subject to much debate.

1.2.1 Inflation

As a general rule, a country with a consistently lower inflation rate exhibit a rising currency value, as its purchasing power increases relative to other currencies. Those countries with higher inflation typically see depreciation in their currency in relation to the currencies of their trading partners. This is also usually accompanied by higher interest rates.

1.2.2 Interest Rates

Interest rates, inflation and exchange rates are all highly correlated. By manipulating interest rates, central banks exert influence over both inflation and exchange rates, and changing interest rates impact inflation and currency values. Higher interest rates offer lenders in an economy a higher return relative to other countries. The impact of higher interest rates is mitigated, however, if inflation in the country is much higher than in others, or if additional factors serve to drive the currency down. The opposite relationship exists for decreasing interest rates - that is, lower interest rates tend to decrease exchange rates.

1.2.3 Current-Account Deficits

The current account is the balance of trade between a country and its trading partners, reflecting all payments between countries for goods, services, interest and dividends. A deficit in the current account shows the country is spending more on foreign trade than it is earning, and that it is borrowing capital from foreign sources to make up the deficit. In other words, the country requires more foreign currency than it receives through sales of exports, and it supplies more of its own currency than foreigners demand for its products. The excess demand for foreign currency lowers the country's exchange rate until domestic goods and services are cheap enough for foreigners, and foreign assets are too expensive to generate sales for domestic interests.

1.2.4 Public Debt

Countries will engage in large-scale deficit financing to pay for public sector projects and governmental funding. While such activity stimulates the domestic economy, nations with large public deficits and debts are less attractive to foreign investors. The reason is that large debt encourages inflation, and if inflation is high, the debt will be serviced and ultimately paid off with cheaper real dollars in the future.

In the worst case scenario, a government may print money to pay part of a large debt, but increasing the money supply inevitably causes inflation. Moreover, if a government is not able to service its deficit through domestic means (selling domestic bonds, increasing the money supply), then it must increase the supply of securities for sale to foreigners, thereby lowering their prices. Finally, a large debt may prove worrisome to foreigners if they believe the country risks defaulting on its obligations. Foreigners will be less willing to own securities denominated in that currency if the risk of default is great. For this reason, the country's debt rating (as determined by Moody's or Standard & Poor's, for example) is a crucial determinant of its exchange rate.

1.2.5 Terms of Trade

A ratio comparing export prices to import prices, the terms of trade is related to current accounts and the balance of payments. If the price of a country's exports rises by a greater rate than that of its imports, its terms of trade have favorably improved. Increasing terms of trade shows greater demand for the country's exports. This, in turn, results in rising revenues from exports, which provides increased demand for the country's currency (and an increase in the currency's value). If the price of exports rises by a smaller rate than that of its imports, the currency's value will decrease in relation to its trading partners.

1.2.6 Political Stability and Economic Performance

Foreign investors inevitably seek out stable countries with strong economic performance in which to invest their capital. A country with such positive attributes will draw investment funds away from other countries perceived to have more political and economic risk. Political turmoil, for example, can cause a loss of confidence in a currency and a movement of capital to the currencies of more stable countries. In Kenya during the December 2007-March 2008 a period when the country experienced the post-election violence, the KES/US$ skyrocketed to the highest levels due to the political instability.

1.3 Characteristics of financial statistics

Financial return statistics exhibit the characteristics called the *stylized facts*. The four facts are listed below (Christian Francq, 2010):

1. The mean of the return series is close to zero.
2. Returns R_t are leptokurtic distribution. The empirically estimated kurtosis is mostly greater than 3
3. The return process is white noise
4. Volatility tends to form clusters. After a large (small) price change (positive or negative) a large (small) price change tends to occur. This effect is called volatility clustering.

Research into the time series models of changing variances and covariance is really important in the daily volatility analysis. Financial series are characterized by clustered periods of large volatility followed by periods that have low volatility. This idea led to the fact that the clustering can be predicted. Auto Regression Conditional Heteroscedacity (ARCH) by Engle (1982) and Generalized Auto Regression Conditional Heteroscedacity (GARCH) by Bollerslev (1986) models are very good at predicting this volatility than the conventional statistical methods. In these models, the key concept is the conditional variance. In the classical GARCH models, the conditional variance is expressed as a linear function of the squared past values of the series. This particular specification is able to capture the main stylized facts characterizing financial series.

Reliable estimates and forecasts are important for large credit institutes where volatility is directly used to measure risk. Smith et al. (1990) found that volatility in prices has implications on the profits and survival of business enterprises as most industries depend on imports and exports. In economics, factors have some forecasting capabilities; the most important factors have been lagged endogenous return. The classics linear AR/ ARIMA processes nor non-linear generalization can fulfill this task, but by use of GARCH model that can replicate stylized facts appropriately (Jurgen Franke, 2008) as it provides an accurate assessment of variances and covariances through its ability to model time-varying conditional variances.

It is with these stylized facts that I want to investigate on the exchange returns since they have a great implications to the Kenyan economy.

Picture 1: Exchange rate chart

CURRENCY	BUY NOTES	SELL NOTES
US DOLLAR $	83.20	83.85
STERLING POUND	137.00	139.00
EURO	122.80	123.65
SWISS FRANCS	95.00	96.00
CANADIAN $	85.85	87.00
JP YEN / 100	85.00	101.90
DANISH KRN	14.40	17.10
NORWEGIAN KRN	13.70	16.45
SWEDISH KRN	11.75	14.50
AUSTRALIAN $	84.50	89.25
INDIAN RUPEE	1.50	2.10
HONKONG $	9.30	11.55
SINGAPORE $	64.00	67.45
SAUDI RIYAL	21.20	22.80
UAE DIRHAM	21.60	23.25
TANZANIA /100	4.50	6.60
SA RAND	10.70	13.40
OMANI RIYAL	207.00	215.95
	Rates as at: 03/05/2011	

This electronic chart picture was taken at Leo Foreign Exchange Bureau (LFB) in Mombasa town on 3rd may 2011 indicating the KES against other international currency rates of exchange.

1.4 Problem Statement

Assume we have observations $Y_1, Y_2, \cdots, Y_t$ of a financial time series of KES/US$ and in this research we are interested in formulating a function of the observed data say, $f(Y_{t-1})$ that would help us model us do predictions, such that the time series can be represented by

$$Y_t = f(Y_{t-1})$$

where $f(Y_{t-1})$ is a smooth function containing an error term function.

From the literature review, financial time series returns exhibit negative skewness and have excess kurtosis such that the function $f(Y_{t-1})$ is heteroscedastic and cannot be modeled assuming normality.

1.5 Research Objectives

Given the function $Y_t = f(Y_{t-1})$, $t \in Z$, the objectives of the research are:

a) To estimate $f(Y_{t-1})$

b) To estimate the parameters of $f(Y_{t-1})$

c) To find the properties of the parameters

d) To perform model diagnostics

e) Test model for forecasting

1.6 Purpose of the research

The main purpose of this research is therefore to forecast the KES/US$ exchange rates return's and volatilities by using time series and GARCH(1,1) approach. By coming up with a mathematical model that can predict the returns in the future, many banks and other stakeholders would strategically know when to trade to maximize their revenues.

1.7 Significance of the research

The research looks at two time series components in the KES/US$ exchange rate; the exchange rate prices and the volatility using the daily returns. By forecasting the returns at a day, $t+k$ in the future, the forex players would be able to predict the returns and hence avert the losses.

1.8 Scope

This study is uses ARIMA-type modelling for the exchange rate and GARCH (p,q) on historical data to measure volatility.

CHAPTER 2

LITERATURE REVIEW

There are several methods used in time series analysis such as Auto regressive process (AR), Moving Average (MA), Auto regressive Moving Average (ARMA) and Auto regressive Integrated Moving Average (ARIMA).

ARIMA models have been used for forecasting different types of time series and have been compared with a benchmark model for its validity. Leseps and Morell (1977) in their study found that the exchange rate follows a long-term trend with short-term fluctuation. Therefore, to capture the long term trend, many authors had used ARIMA model as proposed by Box-Jenkins (1976), to forecast the exchange rate.

Chen (1995) introduced a new pre-differencing transformation for the ARIMA model for forecasting Standard and Poor's (S&P 500) index volatility. The out of sample forecasting performance of the ARIMA model using the pre-differencing transformation was compared with the out of sample forecasting performance of the mean reversion model and the GARCH model. The ARIMA model using the pre-differencing transformation introduced in this study was found to be superior to both the mean reversion model and the GARCH model in forecasting monthly S&P 500 index volatility for the forecast comparison periods used in this study.

Madura, et. al (1999) assessed the forecast bias and accuracy of the three commonly used forecast methods for 12 divergent emerging market currencies. The random walk method outperformed the forward rate and ARIMA methods for some emerging market currencies, and was not outperformed by these alternative methods. In general, it appears that the incorporation of expectation components by the implicit forward and ARIMA methods do not improve the forecast, and actually reduce forecast accuracy in some cases.

Lim and McAleer (2002) used various ARIMA models over the period 1975(1)–1989(4) for tourist arrivals to Australia from Hong Kong, Malaysia and Singapore. The fitted ARIMA model is found to be valid when tourist arrivals were forecasted for Singapore for the period 1990(1)–1996(4).

In the Indian context, Mahadevan (2002) found that while forecasting 10 year government securities yield, ARIMA had a marginally better directional accuracy than that of the moving average model in a static forecast, whereas the lagged moving averages for 10-year government securities outperforms ARIMA model in dynamic forecasting.

The ARCH and GARCH (p,q) model was originally introduced by Bollerslev (1986) and Engle (1982) in order to model the fluctuations of the variances of the time series data. To model returns ARCH and GARCH have been touted to be the best due to heteroscedastic nature of the daily returns. Financial time series have volatility clustering effect and hence conditional heteroscedastic based models are being used to develop a more robust model for forecast. The literature shows that attempt has been made to forecast the returns or the volatility of returns using conditional heteroscedastic-based models. Engle suggests ARCH and GARCH provide better forecasts of variance.

Kearns and Pagan (1991) examined monthly volatility of the Australian stock market over the period 1875-1987, and fitted ARCH and GARCH models to the data. It was found that the GARCH (1, 1) model outperformed the ARCH model for forecasting the volatility of the returns.

Brooks and Lee (1997) used ARCH/GARCH models to investigate Australian financial futures data. The extent to which the parameters of the models change over time, were examined by analyzing the data. The results vary over time and simple models such as the ARCH (1) model provides a reasonably good fit to the data.

Tabak and Guerra (2002) examine the relationship between stock returns and volatility over the period of June 1990 to April 2002. The relationship between stock returns and volatility is tested using seemingly unrelated regressions methods and AR (1)-GARCH (1,1) estimation. They conclude that using both a seemingly unrelated regressions (SUR) methodology and an AR (1)-GARCH (1,1) estimation changes in volatility are negatively related to stock returns.

Radha et. al (2008) applied ARMA, ARIMA-GARCH and Random Walk models on the short-term interest rates forecasting. The results indicate that the short-term interest rates do have volatility clustering effect in the time series and this captured by the GARCH model. Moreover, the ARIMA and random walk model developed were not a good fit. The comparison of the models for forecasting Short-term interest rates - Implicit Yield on 91 day Treasury bill, call money and overnight MIBOR show that ARIMA-GARCH is an appropriate model for forecasting.

Maana et al (2010) applied GARCH (1,1) in the estimation of volatility in the Kenyan foreign exchange market data for the period 1993 - 2006. Exploratory analysis showed that the exchange rates are leptokurtic and slightly positively skewed. This implies that the exchange rate depreciation was preferred during the period, probably to ensure that Kenya's exports remained competitive. The GARCH(1,1) fitted the returns well.

A key role is played by forecasts of exchange rate variability is key for businesses as they reduce uncertainty concerning future changes in prices of individual currencies. Correctly constructed forecasts enable currency market participants to gain proceeds from speculative or arbitrage investments, and are also used in the process of currency risk management, by extension, better measures of Value at Risk (Engle, 2001).

Yu (2002) evaluated the performance of nine alternative models such as the random walk, historical average, Moving average, Simple regression, exponential smoothening, exponentially weighted moving average, ARCH and GARCH. He concludes that the (i) ARCH-type models can perform well or badly depending on the form chosen; (ii) the performance of the GARCH(3,2) model, the best model within the ARCH family, is sensitive to the choice of assessment measures; and (iii) the regression and exponentially weighted moving average models do not perform well according to any assessment measure, in contrast to the results found in various markets.

In the chapter 3 we will explain how to adopt ARIMA-type models to come up with the model for the exchange rates and use GARCH procedure in the estimation of volatility of the KES/US$ exchange rates spanning the period 2nd November 2004 to 31st December 2010.

CHAPTER 3

METHODOLOGY

3.1 Data

To achieve the goal of the study, the data of daily average exchange rate of Kenya Shilling versus US dollar, are taken for 1531 trading days. The daily data is taken from the Central Bank of Kenya's website www.centralbank.go.ke for the period 2nd November 2004 to 31st December 2010, that's from Monday to Friday only on trading days. For simplicity, we will analyze the data as if they were equally spaced.

The choice of the US dollar currency was based on their relative proportions in the Bank's foreign exchange reserve portfolio.. The currency composition of Kenya's imports comprised about 52% in US dollars in December 2006 (CBK, 2006).

3.2 Statistical software

To analyze the data We will use R statistical software. The software is chosen because of its wide application and acceptance in the field of Statistics and Econometrics.

3.3 Model building for the exchange rates

3.3.1 Introduction to time series

In this research We will first graph the trend for the period mentioned. We will then give a description of the results based on the occurrences in Kenya in the period. to model using time series data, the series has to be stationary to give robust results and predictions.

Time series are either stationary or non-stationary, in the case of stationary a time series is stationary if there's no systematic change in the mean i.e. no trend or there's no systematic change in the Variance. A time series is strictly stationary if the joint distribution of $Y_{t_1}, Y_{t_2}, Y_{t_3}, \cdots, Y_{t_n}$ is the same as the joint distribution of $Y_{t_1+h}, Y_{t_2+h}, Y_{t_3+h}, \cdots, Y_{t_n+h}$ and h being the distance between the observations. In other words shifting the time origin by an amount h has no effect on the joint distribution which must only depend on the interval between $t_1, t_2, \ldots, t_n$. That means that for instance if h=1, the distribution of X_t must be the same for all values of t i.e. the mean $\mu_t = \mu$ and $\sigma_t^2 = \sigma_1$ both constants don't depend on t.

To test for stationarity we will use the Autocorrelation coefficient functions (ACF) and Partial Autocorrelation functions (PACF) graphs. If the ACF is decaying exponentially, it implies that the data series is not stationary. However this will be formally tested by using the Augmented Dickey-Fuller (ADF) test for stationarity. ADF tests to determine whether a time series is stationary or, specifically, whether the null hypothesis of a unit root can be rejected.

Consider the model $Y_t = \alpha Y_{t-1} + X_t \quad for \quad t = 1, 2, \ldots$ Where $\{X_t\}$ is a stationary process, the process $\{Y_t\}$ is non-stationary if the coefficient $\alpha = 1$ but it becomes stationary if $abs(\alpha) < 1$. Suppose that $\{X_t\}$ is an AR(k) process: $X_t = \phi_1 X_{t-1} + \cdots + \phi_k X_{t-k} + e_t$. Under the null hypothesis that $\alpha = 1$, $X_t = Y_t - Y_{t-1}$. Letting $a = \alpha - 1$, we have

$$
\begin{aligned}
Y_t - Y_{t-1} &= (\alpha - 1) Y_{t-1} + X_t \\
&= a Y_{t-1} + \phi_1 X_{t-1} + \cdots + \phi_k X_{t-k} + e_t \\
&= a Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2}) + \cdots + \phi_k (Y_{t-k} - Y_{t-k-1}) + e_t
\end{aligned}
$$

Test statistic is basically a t-statistic

$$
t^* = \frac{\hat{a}}{s.e(\hat{a})} \tag{3.1}
$$

Hypothesis tested is

H_0: $a = 0$ (the data is not stationary (unit root))

H_1: $a < 0$ (the data is stationary (no unit root))

Decision rule:

If t* > ADF critical value, we do not reject null hypothesis, i.e., unit root exists.

If t* < ADF critical value, we reject null hypothesis, i.e., unit root does not exist.

This will call for differencing if it is found that $P-\text{value} > \alpha = 0.05$ then we conclude H_0 that the data is not stationary.

This will now call for first-order differencing of logarithms of the data. This is first-order differences are same as calculating the compounded returns or log returns in finance.

This transformation will make the data stationary.

There are several models that are employed in time series analysis, namely: Auto Regressive (AR), Moving Average (MA), Auto Regressive Moving Average (ARMA), and Auto Regressive Integrated Moving Average (ARIMA) among others. We'll discuss them briefly in the following subsections.

3.3.2 Autoregressive model (AR) model

The notation AR (p) refers to the autoregressive model of order p. The AR (p) model is written

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} \ldots \alpha_p X_{t-p} + \varepsilon_t \qquad (3.2)$$

Where ε_t -N (0, σ^2) and i.i.d., α_t are the parameters of the model.

Some constraints are necessary on the values of the parameters of this model in order that the model remains stationary. For example, processes in the AR (1) model with $|\varphi 1| \geq 1$ are not stationary.

The simplest of the AR process is when p=1 i.e. AR (1)

The properties of AR (1) process: $E(Y_t) = 0$

$$Var(Y_t) = \frac{\sigma^2}{1-\theta_1^2} \quad if\ abs(\theta_1) < 1$$

$$Corr(Y_t, Y_{t-k}) = \rho_k \quad if \quad abs(\theta_1) < 1, k > 0$$

3.3.3 Moving Average (MA) model

The notation MA (q) refers to the moving average model of order q:

$$Y_t = \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + .. + \beta_q \varepsilon_{t-q} + \varepsilon_t \quad (3.3)$$

Where β_i 's are the parameters of the model, μ is the expectation of Y_t (often assumed to equal 0), ε_t are white noise error terms and $Var(Y_t) = \sigma^2 \sum_i \beta_i^2$

3.3.4 Autoregressive Moving Average model

The notation ARMA (p,q) refers to the model with p autoregressive terms and q moving average terms. This model contains the $AR(p)$ and $MA(q)$ models,

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_p Y_{t-p} + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q} + \varepsilon_t \quad (3.4)$$

Where $\varepsilon_t \sim N(0,1)$, $p, q \geq 0$ are integers and p, q are the order of the model.

3.3.5 Auto-Regressive Integrated Moving Average (ARIMA)

ARIMA models are, in theory, the most general class of models for forecasting a time series which can be stationarized by transformations such as differencing and logging. In fact, the easiest way to think of ARIMA models is as fine-tuned versions of random-walk and random-trend models: the fine-tuning consists of adding lags of the differenced series and/or lags of the forecast errors to the prediction equation, as needed to remove any last traces of autocorrelation from the forecast errors.

Lags of the differenced series appearing in the forecasting equation are called "auto-regressive" terms, lags of the forecast errors are called "moving average" terms, and a time series which needs to be differenced to be made stationary is said to be an "integrated" version of a stationary series. A non-seasonal ARIMA model is classified as an "ARIMA (p,d,q)" model, where:

- **p** is the number of autoregressive terms,
- **d** is the number of non-seasonal differences, and
- **q** is the number of lagged forecast errors in the prediction equation.

For an $ARIMA(p,d,q)$

is a non-stationary ARMA (p,q). $\Delta^d Y_t = W_t$

$$W_t = \phi_1(W_{t-1}) + \phi_2(W_{t-2}) + \cdots + \phi_p(W_{t-p}) + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_q e_{t-q} \quad (3.5)$$

For $d = 1$

Or in other terms,

$$Y_t - Y_{t-1} = \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \cdots + \phi_p(Y_{t-p} - Y_{t-p-1}) + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \quad (3.6)$$

To identify the appropriate ARIMA model for a time series, we begin by identifying the order(s) of differencing needing to stationarize the series and remove the gross features of seasonality, perhaps in conjunction with a variance-stabilizing transformation such as logging or deflating. If you stop at this point and predict that the differenced series is constant, you have merely fitted a random walk or random trend model. However, the best random walk or random trend model may still have auto correlated errors, suggesting that additional factors of some kind are needed in the prediction equation.

We will develop a multistep model-building strategy espoused so well by Box and Jenkins (1976). There are three main steps in the process, that is; Model specification (or identification), Model fitting, and Model diagnostics.

To determine the order of the exchange rates time series, we follow the steps below;

- Step 1: Examination of time-series plot

  The first step is to produce a time-series plot; namely, to plot X_t against t and examine the plot to identify obvious trends, seasonal components, and outliers. We shall there by conduct tests on the data to ascertain if the series is stationary using the Augmented Dickey Fuller (ADF) test. These components should be removed through differencing or other appropriate methods.

- Step 2. Examination of correlogram

  Trend and seasonal components may show up in a correlogram (i.e., the plot of sample ACF ρ_k against k). A slowly damping correlogram is indicative of a slowly varying trend component. A periodic fluctuating correlogram is indicative of a periodic component (with the same period). Taking the difference at appropriate time lags may remove those non-stationary components.

- Step 3. Determining the MA-order from the ACF and the AR-order from the PACF

  If the data appear stationary in both the time-series plot and correlogram, we may try to identify the order (p,q) from the sample ACF $\{\hat{\rho}(k)\}$ and the sample PACF $\{\hat{\Pi}(k)\}$ first. As a rule of thumb, we fit an AR (p) model to the data if $|\hat{\Pi}_i| \leq 1.96/\sqrt{n}$ for about 95% of k's among all k>p and n is the sample size. Below are the classical characteristics of ACF and PACF of ARMA-type of models.

Table 1: Characteristics of ACF and PACF

Behavior of theoretical ACF and PACF for stationary process		
Model	*ACF*	*PACF*
MA(q)	Cuts off after lag q	Exponential decay and/or damped sinusoid
AR(p)	Exponential decay and/or damped sinusoid	Cuts off after lag p
ARIMA(p,q)	Exponential decay and/or damped sinusoid	Exponential decay and/or damped sinusoid

- Step 4. Determining the orders using AIC, AICC or BIC. Each method has its own merit. A practically relevant question is when to use what, although a general answer to this question is inconceivable. The choice should depend on the nature and the aim of the data analysis. Empirical experience suggests that AIC is a good starting point. AIC is defined as

$$AIC = -2 * \log\text{-}likelihood + 2p \tag{3.7}$$

An alternative widely used criterion is the Bayesian Information Criterion (BIC) which essentially replaces the term $2p$ in the AIC with the expression $p + p\ln N$.

Hence BIC is

$$BIC = -2 * \log\text{-}likelihood + p(1 + \ln N) \tag{3.8}$$

The corrected AIC i.e. AICC

$$AICC = AIC + \frac{2(k+1)(k+2)}{n-k-2} \tag{3.9}$$

Where n is the sample is size and k is the total number of parameters excluding the noise variance.

So far ARIMA-type models concern with the conditional mean structure of time series data however; there are cases where there exist inconsistencies between the data and after basic transformations. For example, squaring the residuals of the data or even taking the absolute values, we expect the ACFs of the data to remain insignificant but any deviations from this would imply the data exhibits high order dependency in its structure. To counter this we'll test for conditional heteroscedacity using McLeod and Li test of the squared residuals fitting the ARIMA-type model after the transformations, if the test is significant we reject the absence of conditional heteroscedacity.

3.4 Model building for the returns

Returns are defined by

$$r_t = \left(\frac{Y_t - Y_{t-1}}{Y_t} \right) \tag{3.10}$$

Where {Y} are the exchange rates at the respective times.

Now, if we have many periods, say, N, in one time interval (i.e. 12 months in an interval of one year) then the total return of the entire interval is simply the sum of all the individual period's returns, i.e.

$$r_t = \sum_{t=1}^{N} r_t = \sum_{t=1}^{N} \left(\frac{Y_t - Y_{t-1}}{Y_{t-1}} \right) \tag{3.11}$$

However, what happens if we start to slice time into smaller and smaller intervals. Say, we start slicing time (one year) into minutes, seconds, nanoseconds and so on until we get to the mathematical definition of an infinitesimally small interval of time. We are now talking about the limit when delta t (the smallest measurable unit of time) goes to zero. Mathematically speaking, we say $\Delta t \to 0$. In the limit, the above expression for return will reduce to:

$$r_t = \lim_{\Delta t \to 0} \sum_{t=1}^{N} r_t = \lim_{\Delta t \to 0} \sum_{t=1}^{N} \frac{\Delta Y}{Y}$$

In the limiting case if $\Delta t \to 0$, then $\Delta Y \to dY$ and the summation sign will get replaced by an integral. Therefore, the expression for the return (dropping the subscript for time) becomes:

$$R_t = \ln\left(\frac{Y_t}{Y_{t-1}}\right) \tag{3.12}$$

In this research therefore we will model the returns as indicated in equation 3.12.

Conditional heteroscedastic models are the econometric tools used to estimate and forecast asset returns volatility. In this research we will consider the application of the GARCH (1, 1) process in the estimation of volatility in the exchange rates. We will apply a quasi-maximum likelihood estimation (QMLE) procedure and will come up with asymptotic properties of its estimators. We will carry out exploratory data analysis to the data to come up with the underlying structure of the data series.

The GARCH (1, 1) is the most robust of the family of volatility models. GARCH models have been applied to a wide range of time series analyses, but applications in finance have been particularly successful and have been the focus of this introduction.

3.4.1 GARCH

In that case, the GARCH(p,q) model (where p is the order of the GARCH terms σ^2 and q is the order of the ARCH terms R^2) is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2 + \cdots + \alpha_q R_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \cdots + \beta_p \sigma_{t-p}^2 \tag{3.13}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i R_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{3.14}$$

The lag length p of a $GARCH(p,q)$ process is established in three steps:

1. Estimate the best fitting AR(q) model

$$Y_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i Y_{t-i} + \varepsilon_t \tag{3.15}$$

2. Compute and plot the autocorrelations of ε_t^2 by

$$\rho = \frac{\sum_{t=i+1}^{T}\left(\hat{\varepsilon}_t^2 - \hat{\sigma}_t^2\right)\left(\hat{\varepsilon}_{t-1}^2 - \hat{\sigma}_{t-1}^2\right)}{\sum_{t=1}^{T}\left(\hat{\varepsilon}_t^2 - \hat{\sigma}_t^2\right)^2} \tag{3.16}$$

3. The asymptotic that is for large samples standard deviation of $\rho(i)$ is $1/\sqrt{n}$. Individual values that are larger than this indicate GARCH errors. To estimate the total number of lags, use the *Ljung-Box test* until the value of these are less than. say, 10% significant. *The Ljung-Box Q-statistic* follows χ^2 distribution with n degrees of freedom if the squared residuals ε_t^2 are uncorrelated. It is recommended to consider up to T/4 values of n. The null hypothesis states that there are no ARCH or GARCH errors. Rejecting the null thus means that there exists such error in the conditional Variance.

3.4.2 Estimating of GARCH parameters by Quasi Maximum Likelihood

A GARCH, model with orders $p \geq 1$ and $q \geq 0$ is defined as

$$R_t = \sigma_t z_t \tag{3.17}$$

and

$$\sigma_t^2 \equiv \sigma_t(\theta)^2 = c + \sum_{i=1}^{p} b_i R_{t-i}^2 + \sum_{j=1}^{q} a_j \sigma_{t-j}^2 \tag{3.18}$$

where c>0, $b_i \geq 0$ and $a_j \geq 0$ are unknown parameters, $\theta = (c, b_1, b_2 ..., b_p, a_1, ..., a_q)$, $z_t \sim N(0,1)$ and z_t is independent of $\{X_{t-k}, k \geq 1, \quad \forall t\}$ (Christian Francq, 2010). θ belongs to a parameter space of the form $\Theta \subset (0, +\infty) * (0, \infty)^{p+q}$

The distribution of R_t is unknown. When $q = 0$, reduces to an ARCH model. The necessary and sufficient condition for above equation to define a unique strictly stationary process {X_t = 0, ±1, ±2,· · ·} with $E(X_t^2)<\infty$ is that $\sum_{i=1}^{p} b_i + \sum_{j=1}^{q} a_j < 1$

Where $E(X_t)=0$ and $Var(X_t)=c/\left(1-\sum_{i=1}^{p} b_i + \sum_{j=1}^{q} a_j\right)$

Under the condition $\sigma_t^2 \equiv \sigma_t(\theta)^2$ may be expressed as

$$\sigma_t(\theta)^2 = \frac{c}{1-\sum_{j=1}^{q} a_j} + \sum_{i=1}^{p} b_i R_{t-i}^2 + \sum_{i=1}^{p} b_i \sum_{k=1}^{\infty} \sum_{j_1}^{q} \cdots \sum_{j_k}^{q} a_{j_1} \cdots a_{j_k} R_{t-i-j_1-\cdots-j_k}^2 \tag{3.19}$$

Where the multiple sums vanishes if $q = 0$.

The true value of the parameter θ is unknown and is denoted by

$\theta_0 = (c, a_1, a_2, \cdots, a_q, b_1, b_{2,\ldots} b_p)$

To write the likelihood of the model, a distribution must be specified for the i.i.d. variables η_t. Here we don't make any assumption on the distribution of these variables, but we work with a function called the (Gaussian) quasi-likelihood which conditionally on some initial values coincides with the likelihood when the η_t are distributed as standard Gaussian. Given initial values $\varepsilon_0, \varepsilon_1, \cdots, \varepsilon_{1-q}, \delta_0^2, \delta_1^2 \cdots \delta_{1-p}^2$ to be specified below, the conditional Gaussian quasi-likelihood is given by

$$L_n(\theta) = L_n(\theta; \varepsilon_1, \varepsilon_2, \cdots \varepsilon_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\Pi\sigma_t^2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_t^2}\right) \tag{3.20}$$

where δ_t^2 are recursively defined for $t \geq 1$ by

$$\tilde{\sigma}_t^2 \equiv \tilde{\sigma}_t(\theta)^2 = c + \sum_{i=1}^{p} b_i X_{t-i}^2 + \sum_{j=1}^{q} a_j \sigma_{t-j}^2 \tag{3.21}$$

For a given value of θ, under the second-order stationarity assumption, the unconditional variance corresponding to θ is a reasonable choice for the unknown initial values:

$$R_0^2 = \cdots = R_{1-q}^2 = \sigma_0^2 = \cdots = \sigma_{1-p}^2 = c \quad (3.22)$$

A QMLE is defined as measurable solution of the equation

$$\phi_n = \arg\min_{\phi\in\Theta} \tilde{I}_n(\phi) \quad (3.23)$$

Taking the logarithm, it is seen that maximizing the likelihood is equivalent to minimizing with respect to θ,

$$\tilde{I}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\tilde{I}_t \quad (3.24)$$

$$\text{Where } \tilde{I}_t = \tilde{I}(\theta) = \frac{R_t^2}{\sigma_t^2} + \log\sigma_t^2$$

The choice of the initial values is unimportant for the asymptotic properties of the QMLE, however in practice this choice may be important.

Note that other methods are possible for generating the sequence σ_t^2 ; for example, by taking $\sigma_t^2 = c_0(\theta) + \sum_{i=1}^{t-1} c_i(\theta) R_{t-i}^2$ where the $c_i(\theta)$ are recursively computed.

Note that for computing $I_n(\theta)$, this procedure involves a number of operations of order n^2, whereas the one we propose involves a number of order n. It will be convenient to approximate the sequence $\tilde{I}_t(\theta)$ by an ergodic stationary sequence. Assuming that the roots of $\beta_0(z)$ are outside the unit disk, the non-anticipative and ergodic strictly stationary sequence $(\sigma_t^2)_t = \{\sigma_t^2(\theta)\}_t$ is defined as the solution of

$$\sigma_t^2 = c + \sum_{i=1}^{p} b_i R_{t-i}^2 + \sum_{j=1}^{q} a_j \sigma_{t-j}^2, \quad \forall t \in Z^+ \qquad (3.25)$$

Where Z^+ is a set of positive integers.

3.4.3 Model identification

As defined in equations 3.7,3.8 and 3.9 AIC, BIC and AICC are used to select the appropriate model. However, empirical evidence, according to Bollerslev et al. (1992), has shown that whilst relatively long lags are required in ARCH models, the GARCH (1, 1) is usually adequate in describing many financial time series. This paper has adopted the latter one. So we select the model with the least combination of the Information Criterion statistic.

To test for dependence in the data we use the sample auto-correlation function. Consider any sequence of data $Y_1, Y_2, ..., Y_n$ then the autocorrelation function r_k for a variety of lags $k = 1,2,...$ is calculated as

$$\rho_k = \frac{\sum_{t=k+1}^{n} \left(Y_t - \bar{Y}\right)\left(Y_{t-k} - \bar{Y}\right)}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2} \qquad (3.26)$$

for $k = 1,2,...$

So we a plot of ρ_k versus k is often called the auto-correlogram. However, ACF is not as useful in the identification of the order of an AR (p) process for which it will most likely have a mixture of exponential decay and damped sinusoid expressions.

Partial Autocorrelation Function (PACF)

The PACF between y_t and y_{t-k} is the autocorrelation between y_t and y_{t-k} after adjusting for $y_{t-1}, y_{t-2}, \cdots, y_{t-k+1}$.

$$Corr(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \cdots, Y_{t-k+1}) = \varphi_{kk} \tag{3.27}$$

Where

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}$$

$$where \quad \phi_{kj} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j}, \; j = 1,2,\cdots,k-1.$$

3.4.4 Tests to be conducted

Testing for autocorrelation-Ljung-Box test

Ljung-Box test is performed to test whether series have significant autocorrelation or not;

The hypothesis is;

H_0: data are not correlated.

H_1: data are correlated (not random).

Ljung-Box test statistic for a number of tested lags k is

$$Q(k) = N(N+2)\sum_{i=1}^{k} \frac{\hat{\rho}_i^2}{N-i} \tag{3.28}$$

where N is the sample size, $\hat{\rho}_i^2$ is the sample autocorrelation at lag k.

Null hypothesis is rejected at $\alpha\%$ significance level if $Q(k) > \chi^2_{1-\alpha,k}$, where $\chi^2_{1-\alpha,k}$ is a $\alpha-$quantile of the chi-square distribution with k degrees of freedom

Testing for Normality of the data

Some of the stylized facts are determined by measuring the third and fourth moment's i.e. skewness and kurtosis respectively.

- Skewness is the ratio of the third order moment and is defined as

$$S = \frac{E(Y-\mu_Y)^3}{\sigma_Y^3} \tag{3.29}$$

- For normality assumption, $S=0$, and kurtosis is a ratio of the fourth order moment, which is assumed to exist to the squared second-order moment and is represented as

$$K = \frac{E(Y-\mu_Y)^4}{\sigma_Y^4} \tag{3.30}$$

And under normality assumption, $k=0$

- QQ plots: It tests for normality and the randomness of the data, we use the QQ plots are used. Normality can be checked more carefully by plotting the so-called normal scores or QQ plot. Such a plot displays the quantiles of the data versus the theoretical quantiles of a normal distribution. With normally distributed data, the QQ plot looks approximately like a straight line and hence the values are expected to fall along or close to the line otherwise the data isn't.

- Shapiro-Wilk's test: It essentially calculates the correlation between the residuals and the corresponding normal quantiles. The lower this correlation, the more evidence we have against normality alternatively, if $p-value > \alpha$ we reject the null hypothesis.

Conditional heteroscedacity effects test

To test for ARCH effects, we'll conduct a

- McLeod and Li test of conditional heteroscedacity of the data, the test statistic is given by

$$Q = N(N+2)\sum_{k=1}^{L}\frac{\hat{\rho}_k^2(R^2)}{N-k} \tag{3.31}$$

- Box-Ljung test with the squared residuals, the test statistic is given by

$$Q=N(N+2)\sum_{k=1}^{L}\frac{\hat{\rho}_k^2(\varepsilon)}{N-k} \tag{3.32}$$

Where N= sample size, L= the number of autocorrelations included in the statistic, and $\hat{r}_k^2$ is the squared sample autocorrelation of residual series $\{\varepsilon_t\}$ at lag k. Under the null hypothesis of model adequacy, the test statistic is asymptotically $\chi^2(L-p-q)$ distributed.

In this the null hypothesis is the absence of the conditional heteroscedacity. If $p-value<0.05$ we reject the null hypothesis and conclude the presence of the effects.

3.4.5 Forecasting using GARCH (1,1)

Predictions using GARCH (1, 1) model can be made by repeated substitutions. First, we provide an estimate for the expected squared residuals:

Given that $R_t=\ln\left(\frac{Y_t}{Y_{t-1}}\right)$ and that

$R_t=\sigma_t z_t;\qquad z_t\sim N(0,1)$ and σ_t is the standard deviation of the data.

More so

$$\left.\begin{aligned}E(R_t)&=0\\ Var(R_t)&=\sigma_t^2\end{aligned}\right\} \tag{3.33}$$

And we know that

$$\sigma_t^2=\alpha_0+\alpha_1 R_t^2+\beta_1\sigma_{t-1}^2 \tag{3.34}$$

And we assume that the probability distribution function (pdf) of R_t is approximately normal

$$f(R_t)=\frac{1}{\sqrt{2\Pi\sigma_t^2}}\exp-\frac{1}{2}\left(\frac{R_t-E(R_t)}{\sigma_t}\right)^2 \tag{3.35}$$

We can say

$$E\left[\sigma_t^2\right] = E\left(\alpha_0 + \alpha_1 R_t^2 + \beta_1 \sigma_{t-1}^2\right)$$
$$= \alpha_0 + \alpha_1 E(R_t^2) + \beta_1(\sigma_{t-1}^2) \qquad (3.36)$$

This implies

$$E(\sigma_t^2) = \alpha_0 + (\alpha_1 + \beta_1)E(\sigma_{t-1}^2)$$
$$\alpha_0 = \left[1 - (\alpha_1 + \beta_1)\right] E(\sigma_{t-1}^2)$$

We can say that the long run variance of a GARCH(1,1) converges to

$$E\left(\sigma_t^2\right) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \qquad (3.37)$$

The conditional variance σ_t^2 and 1-step ahead forecast is known at time τ :

Using the fact that $E\left[\sigma_{t+1}^2\right] = \sigma_{t+1}$ we obtain

$$\sigma_{t+2}^2 = \alpha_0 + \alpha_1 R_{t+1}^2 + \beta_1 \sigma_{t+1}^2 = \alpha_0 + (\alpha_1 + \beta_1)\sigma_{t+1}^2$$

Similarly,

$$\hat{\sigma}_{t+2}^2 = \alpha_0 + (\alpha_1 + \beta_1)\sigma_{t+1}^2 = \alpha_0 + \alpha_0(\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 \sigma_{t+1}^2$$
$$= \alpha_0 + \alpha_0(\alpha_1 + \beta_1) + \alpha_0(\alpha_1 + \beta_1)^2 + (\alpha_1 + \beta_1)^2\left[\alpha_1 R_t^2 + \beta_1 \sigma_t^2\right]$$

Therefore, considering forecasting horizon τ , we have

$$\sigma_{t+\tau}^2 = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)} + (\alpha_1 + \beta_1)^\tau \left[\alpha_1 R_t^2 + \beta_1 \sigma_t^2\right] \qquad (3.38)$$

Moreover, if $\alpha_1 + \beta_1 < 1$ the forecast will converge to the unconditional variance as indicated in equation 3.37 above:

$$\hat{\sigma}_t^2 \to \frac{\alpha_0}{1-(\alpha_1+\beta_1)} \tag{3.39}$$

The same reasoning may be applied for GARCH models of higher orders allowing us to compute multistep ahead forecasts.

In the next chapter we will analyze the data based on this methodology.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Exploratory data analysis

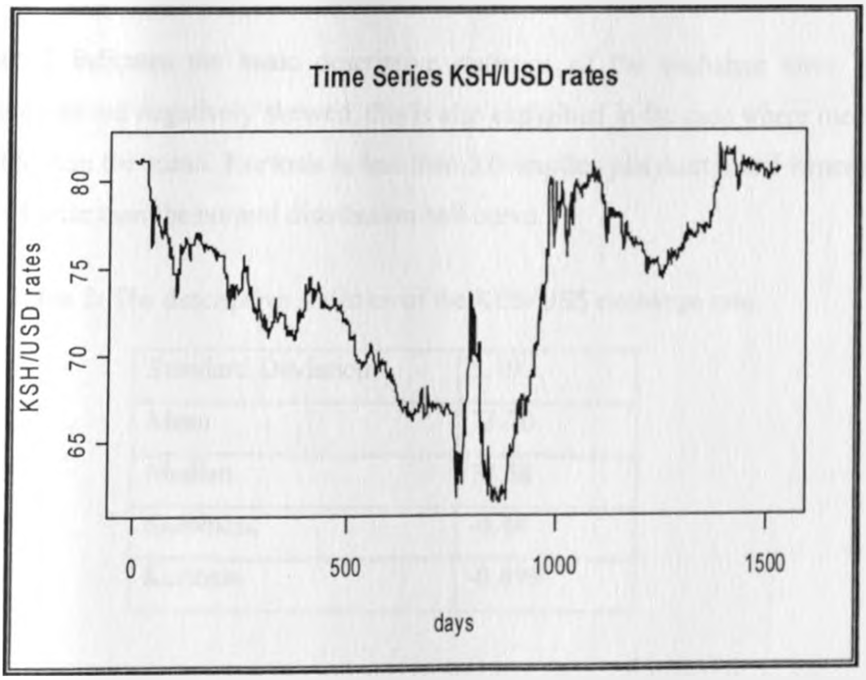
The exchange rates indicated are represented by the following relationship

$$\frac{US\$}{KSh} = \lambda \qquad (4.1)$$

Where λ is the daily average exchange rate.

Because of the inverse relation, if the exchange rate increases it implies the KES has depreciated and vice versa. The trend of the exchange rate is shown in figure 1 below.

Figure 1: A time plot the daily average KES/US$ exchange rates



The time plot easily shows that there was a general strengthening Kenya shilling from Jan 2005 to Dec 2007. This was attributed to the socio-economic reforms and transformation as a result of the NARC government headed by President Mwai Kibaki which came to power in the year 2002 on the platform of economic growth and zero-tolerance on corruption which made the major development partners and donors have great confidence in the country. Indeed the period 2003-2007 witnessed high economic growth rates in the country. The period from around January 2008 shows a weakening of the Kenya shilling that is seen with in the weakening of the KES in the rates in the first 3months of 2008. indeed this was the period when the country faced post election violence that paralyzed the country's economic activities after the disputed 2007 general elections. Once the coalition government was formed in the 1st quarter of 2008, the KES strengthened again.

There was weakening of the KES from May 2008, indeed this period was when the world faced the global financial crisis that highly affected the financial sectors in the USA, UK and Germany and many other countries around the world. Thereafter there is a steady weakening KES up to around January 2009.from about mid-2009, the KES strengthened up to late 2009 when it weakened till January 2011.

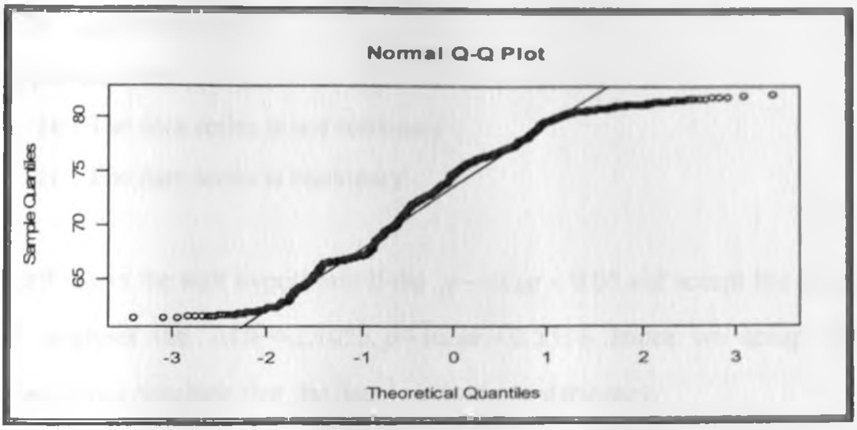
In table 2 indicates the basic descriptive statistics of the exchange rates. The exchange rate are negatively skewed, this is also explained in the case where median is greater than the mean. Kurtosis is less than 3.0 implies platykurtic and hence the curve is flatter than the normal distribution bell curve.

Table 2: The descriptive statistics of the KES/US$ exchange rate

Standard Deviation	5.19
Mean	73.70
Median	74.68
Skewness	-0.44
Kurtosis	-0.699

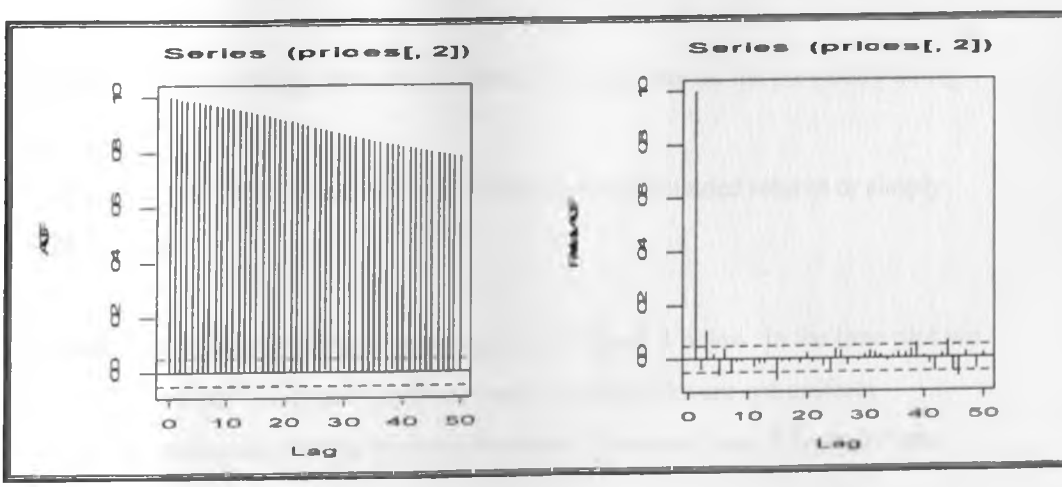
The QQ plot indicates that indeed the rates are not normally distributed as emphasized in the Shapiro test for normality.

Figure 2: QQ plot of the exchange rates



The figure 3 below indicates that the ACF are decaying exponentially, this indicates that the rates are not stationary. For us to do any model fitting on a time series we need to stationarise the rates. In this case we difference to gain stationarity.

Figure 3: ACF and PACF for the rates



The appropriate AR and MA terms are identified for each of the exchange rate using the ACF and PACF, which indicates the significant lags for the MA and AR terms. Looking at the daily exchange rates means, the ACF is decaying exponentially and the PACF exhibit damped sinusoid. Using this empirical analysis, the prices exhibit AR(p) and MA(q). We shall formally therefore conduct ADF test for stationarity $\alpha = 0.05$ significance level.

The hypotheses are;

H_0: The data series is not stationary

H_1: The data series is stationary

We shall reject the null hypothesis if the $p-value < 0.05$ and accept the alternative. In the analysis the ADF=-2.1055, $p-value = 0.5336$ hence we accept the null hypothesis and conclude that the data is indeed not stationary.

This calls for a transformation that would make the data stationary; in this case we take first order difference of the natural logarithms of the exchange rates and is determined as

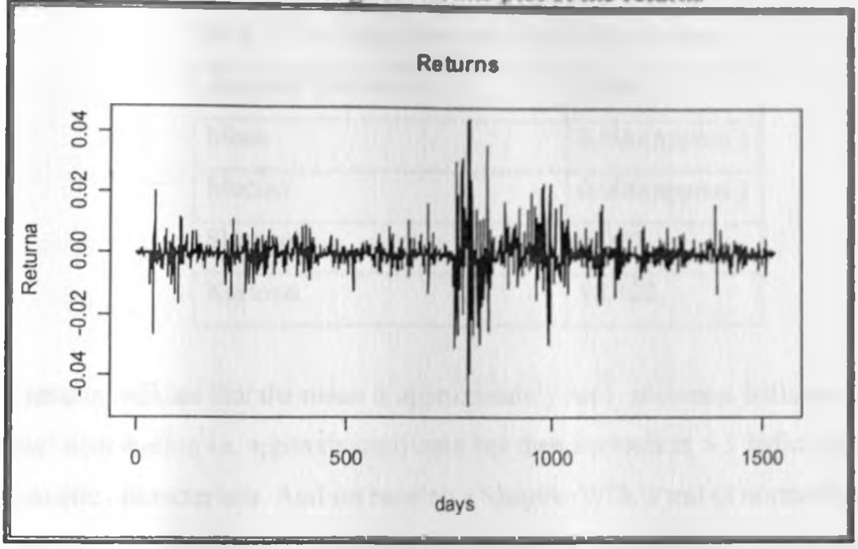
$$R_t = \ln\left(\frac{Y_t}{Y_{t-1}}\right) \tag{4.2}$$

Where Y_t is the exchange rate on day t and Y_{t-1} is the rate on the previous trading day $t-1$.

The equation 4.2 also is used in the calculation of compounded returns or simply called returns.

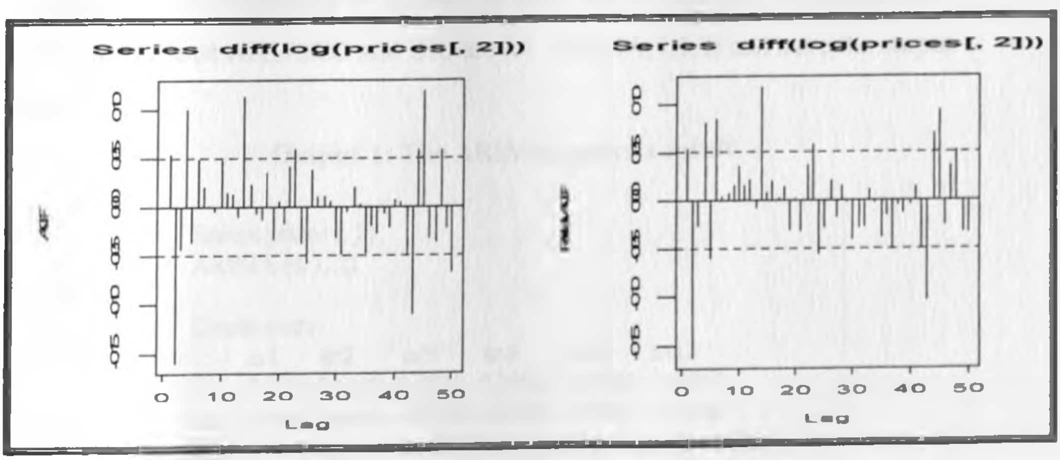
This transformation gives the time plot output in figure 4 below. In the time plot we now see the volatility clustering phenomena i.e. the shocks are not uniform throughout, hence introducing the heteroscedastic characteristic of financial time series.

Figure 4: Time plot of the returns



We can conduct an ADF test again to check if the log-differencing has made the series stationary. The test indeed indicates $ADF = -10.2248$, $p-\text{value} < 0.05$ implying the time series has become stationary. We can confirm this ADF test by plotting the ACF and PACF of the log-differenced data. The ACF stabilizes, indeed after the differencing the ACF and PACF on differencing once are shown in figure 5 below indicating that the data tends to stabilize after the single differencing of the logs.

Figure 5: ACF and PACF of the returns



We then investigate the descriptive statistics, and are shown in table 3 below.

Table 3: The descriptive statistics of the returns

Standard Deviation	0.006
Mean	0.000(approx.)
Median	0.000(approx.)
Skewness	0.077
Kurtosis	18.122

The results indicate that the mean is approximately zero, skewness indicates a near-normal distribution i.e. approximately zero but then kurtosis is >3 indicating leptokurtic characteristic. And on running a Shapiro-Wilk's test of normality indeed we get $W = 0.7609, p < 0.05$ and hence conclude that the returns are not normally distributed. All these have been identified in literature review to be the main characteristics of financial time series data returns.

Once we have established that the data is now stationary, we can now start the process of identifying the parameters, we will attempt to fit an ARIMA model.

4.2 Exchange Rates Model Estimation

Using several combinations of ARIMA model's for the analysis, the parameters that have the least AIC,AICC and BIC are for ARIMA(4,1,2) and have the output 1 below.

Output 1: The ARIMA analysis result

```
Series: prices [, 2]
ARIMA (4, 1, 2)

Coefficients:
        ar1    ar2     ar3     ar4     ma1     ma2
     0.0815 0.6229 -0.0776  0.2096 -0.0146 -0.7862
s.e. 0.0686 0.0666  0.0250  0.0252  0.0669  0.0645
Sigma^2 estimated as 2.898e-05: log likelihood = 5822.33
AIC = -11630.66 AICc = -11630.58   BIC = -11593.33
```

And as ARIMA model defined in equation 3.5, the equation now becomes.

$$R_t = 0.0815R_{t-1} + 0.6229R_{t-2} - 0.0776R_{t-3} - 0.2096R_{t-4} + e_t + 0.0146e_{t-1} + 0.7862e_{t-2} \quad (4.3)$$

Before using the model for forecasting, it must be checked for adequacy. On doing the diagnostics, the figure 6 below shows that the ACF and PACF of the residuals are generally not significant.

We'll formally test for autocorrelation as a group by applying the Ljung-Box test. The hypothesis are;

H_0: The residuals are uncorrelated

H1: The residuals are correlated

We get the test statistic Q=0.001, p-value=0.9745>0.05, hence we accept that the residuals are uncorrelated.

A model is adequate if the residuals left over after fitting the model are simply white noise. We can now say that it's the best fitting model. The standardized residuals however indicate volatility clustering, that is, periods of low volatilities being followed by periods low-volatilities and vice versa.

If time series values are truly independent, then nonlinear instantaneous transformations such as taking absolute values or squaring preserve independence. However, the same is not true of correlation, as correlation is only a measure of linear dependence. Higher-order serial dependence structure in data can be explored by studying the autocorrelation structure of the absolute returns (of lesser sampling variability with less mathematical tractability) or that of the squared returns (of greater sampling variability but with more manageability in terms of statistical theory). If the returns are independently and identically distributed, then so are the absolute returns (as are the squared returns), and hence they will be white noise as well. The Variance is 0.00004 and mean -0.000003, the mean is not statistically significantly different from zero. However, the volatility clustering observed in the return data (see figure 6 (standardized residuals)) gives us a hint that they may not be i.i.d. otherwise the variance would be constant over time.

Figure 6: The diagnostics of the residuals returns

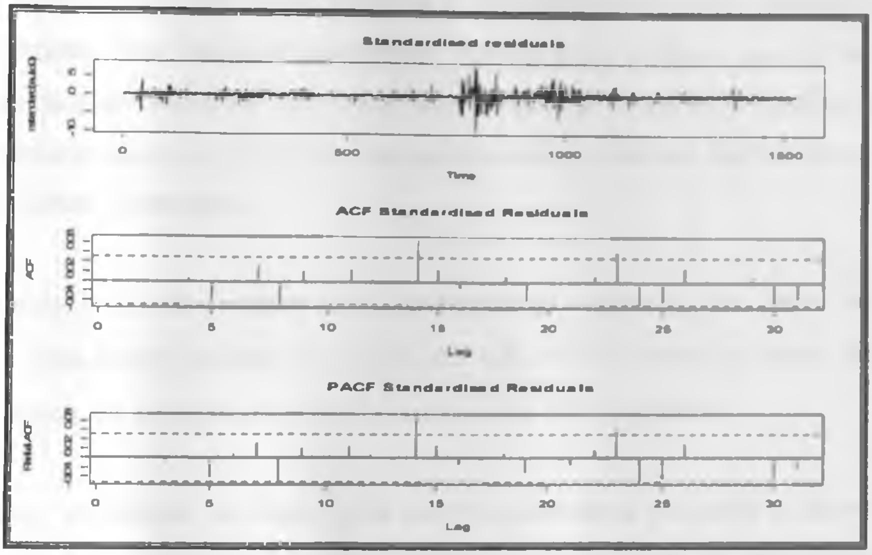
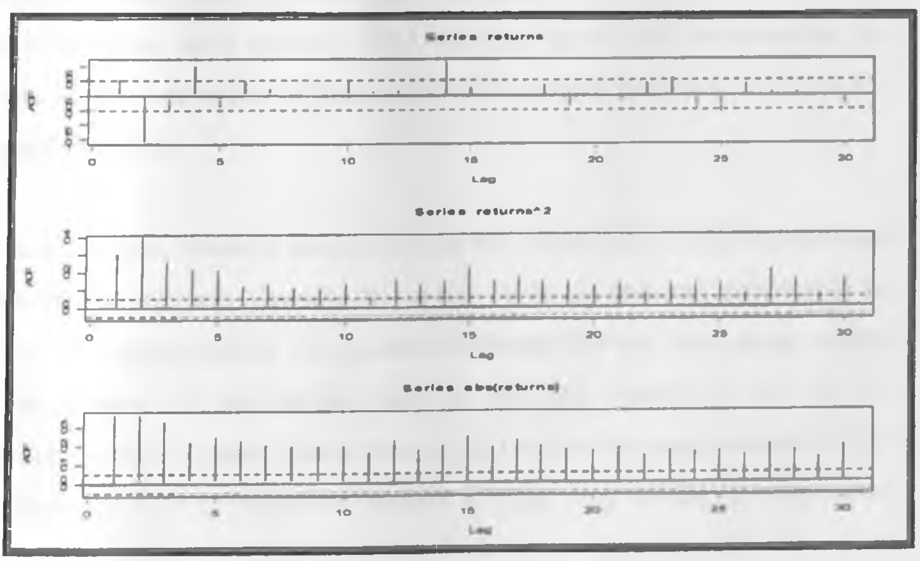


Figure 7: Comparison of ACFs after several transformations



In figure 7, the ACFs show several characteristics, which have several inconsistencies. That is, if we take the squares of the returns and also take absolute values, the ACFs become significant. All the returns with the various transformations are supposed to indicate white noise characteristics, but that is not the case here! Hence, if the transformations of the returns admit some significant autocorrelations, then these autocorrelations furnish some evidence against the hypothesis that the returns are i.i.d. These plots indicate the existence of significant autocorrelation patterns in the absolute and squared data and indicate that the returns are in fact serially dependent.

Further analysis on the residuals using Shapiro-Wilk's normality test under the normality assumption indicate $W = 0.7785, p-value < 0.05$ hence we reject the hypothesis that the residuals are normally distributed at 95% confidence

In summary, the returns are found to be serially uncorrelated but admit a higher-order dependence structure, namely volatility clustering, and a heavy-tailed distribution.

4.3 Testing for Conditional Heteroscedacity

So far ARIMA-type models concern with the conditional mean structure of time series data however, more recently, there has been much work on modeling the conditional variance structure of time series data—mainly motivated by the needs for financial modeling.

These visual tools are formally testing whether the squared data are auto-correlated using the Box-Ljung test. Because no model fitting is required, the degree of freedom of the approximating Chi-square distribution for the Box-Ljung statistic equals the number of correlations used in the test. Hence, if we use m autocorrelations of the squared data in the test, the test statistic is approximately Chi-square distributed with m degrees of freedom, if there is no ARCH. In which case, the first m autocorrelations of the squared residuals from this model can be used to test for the presence of ARCH.

The hypothesis are;

H_0: No ARCH effect in residual squared

H_1: Presence of ARCH effect in residual squared

We get that $\chi^2_{(1)} = 162.05, p-value < 0.05$, we conclude the alternative hypothesis, that there's presence of ARCH effects to the data.

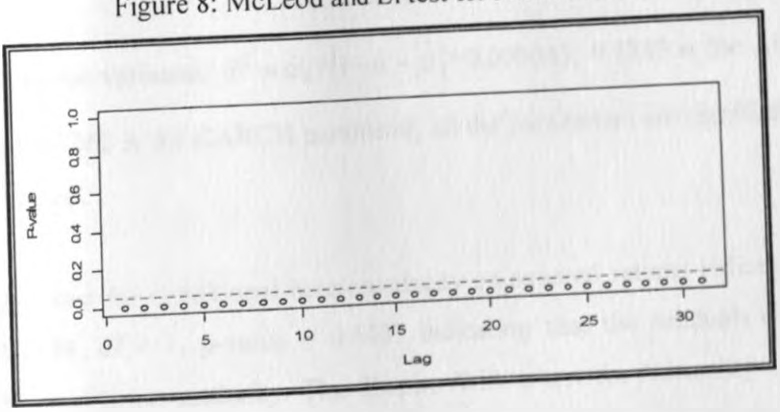
This can be confirmed by applying the McLeod and Li test which checks for the presence of conditional heteroscedacity by computing the Ljung-Box (portmanteau) test with the squared residuals from an ARIMA model.

The McLeod and Li test statistic is given by

$$Q = N(N+2)\sum_{k=1}^{L} \frac{\hat{r}_k^2(\varepsilon^2)}{N-k} \quad (4.4)$$

The McLeod and Li test output shown in the figure below it indicates that all the P-values are significant indicating a strong ARCH effect in the daily returns.

Figure 8: McLeod and Li test for ARCH effects



Having realized that the returns have conditional heteroscedacity effects we will now fit an ARCH-type model. It is commonly observed that such characteristics are rather prevalent among financial time series data. The GARCH models introduced in the next sections attempts to provide a framework for modeling and analyzing time series that display such characteristics.

4.4 GARCH

On doing GARCH analysis we get the following output.

Output 2: The GARCH (1,1) fit output

```
Coefficient(s):
Estimate  Std. Error  t value Pr(>|t|)
a0 8.050e-07  6.283e-08   12.81  <2e-16 ***
a1 1.849e-01  1.372e-02   13.48  <2e-16 ***
b1 7.962e-01  1.115e-02   71.42  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Diagnostic Tests:
Box-Ljung test
data:  Squared.Residuals
X-squared = 0.2179, df = 1, p-value = 0.6407
```

Interpretation:

The GARCH (1,1) model is

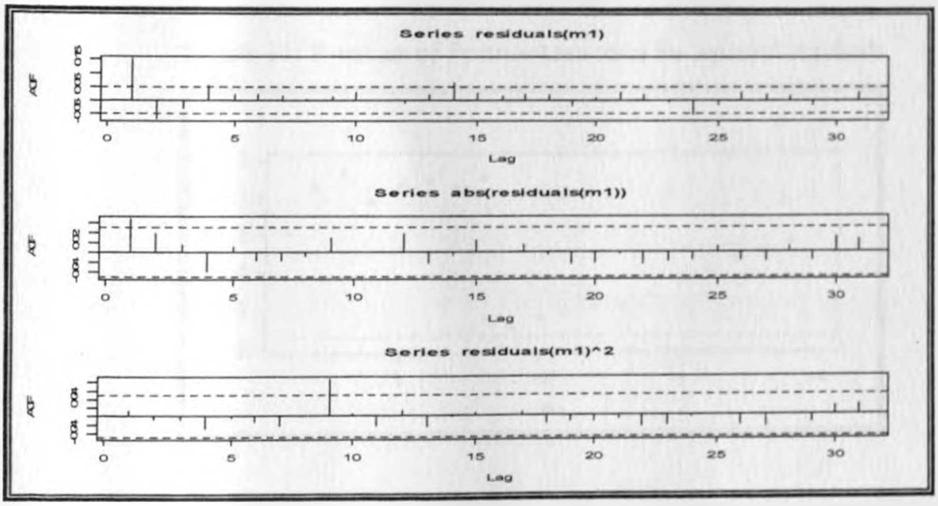
$$\sigma_t^2 = 0.00000081 + 0.1849 r_{t-1}^2 + 0.7962 \sigma_{t-1}^2 \qquad (4.5)$$

Where the long-run variance, $\hat{\sigma}^2 = \hat{\alpha}_0 / \left(1 - \hat{\alpha} - \hat{\beta}\right)$=0.000043, 0.1849 is the ARCH parameter and 0.7962 is the GARCH parameter, all the parameters are significant at almost 99.9% level.

The Ljung-Box test for conditional heteroscedacity on squared returns indicates X-squared = 0.2179, df = 1, p-value = 0.6407 indicating that the residuals do not exhibit conditional heteroscedacity. The Shapiro-Wilk's test for Normality on the residuals gives $W = 0.9037, P-value < 0.05$ under the null hypothesis that the residuals are normally distributed; we therefore conclude that the residuals are not normally distributed.

If the GARCH model is correctly specified, then the standardized residuals should be close to i.i.d. by examining their ACF. Figure 12 indicate the sample ACF of the returns, squared and absolute standardized residuals from the fitted GARCH (1,1) model. The (individual) critical limits in the figure are based on the $1/n$ nominal variance under the assumption of i.i.d. data. This nominal value could be very different from the actual variance of the autocorrelations of the squared residuals even when the model is correctly specified. Nonetheless, the general impression from the figure 9 is that the squared and absolute residuals are serially uncorrelated.

Figure 9: ACF of the residuals, Absolute and Squared residuals of GARCH (1,1)



This could also be reinforced by conducting Ljung-Box test (also called portmanteau test) which Perform a goodness-of-fit test for the GARCH model by checking whether the standardized residuals are i.i.d. based on the ACF of the absolute residuals or squared residuals (Chan, 2008). Figure 10 and 11 displays the p-values of the generalized portmanteau tests with the absolute and squared standardized residuals, respectively, from the fitted GARCH(1,1) model of the exchange rate data for m = 1 to 20. All p-values are higher than 5%, suggesting that the squared residuals are uncorrelated over time, and hence the standardized residuals are independent and indicates that the model fits the data well.

Figure 10: P-values of Portmanteau tests for absolute residuals

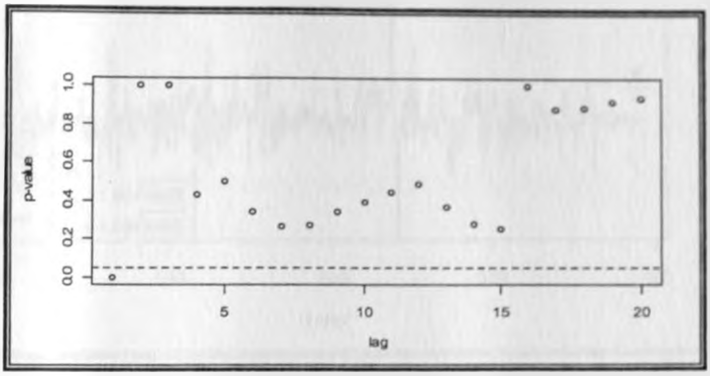
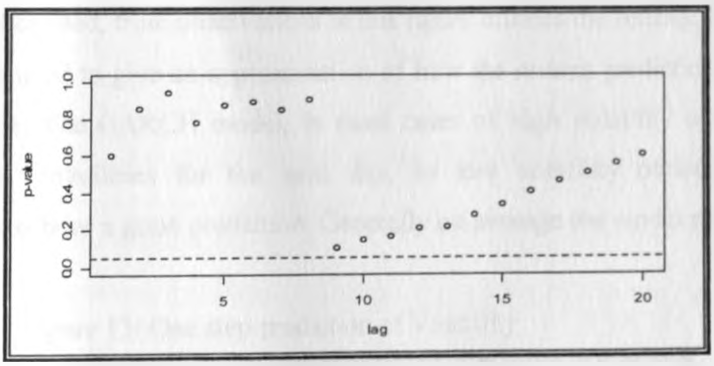


Figure 11: P-values of Portmanteau tests for squared residuals



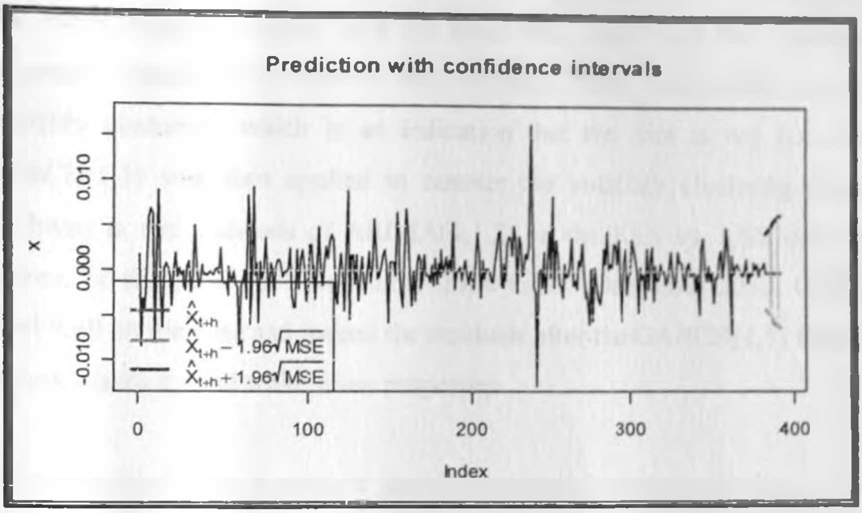
Having noticed that there are no autocorrelations in the residuals, we can now predict for a time τ in future.

4.5 Prediction of Volatilities

Given that the GARCH (1,1) model provides a good fit to the exchange rate data, we may use it to forecast the future conditional variances. We can predict at 95% confidence for example for a 10 days trading ahead. This can be seen in the figure 13 below indicates the forecasted period (note that the index axis is $n*x = .25*1531$

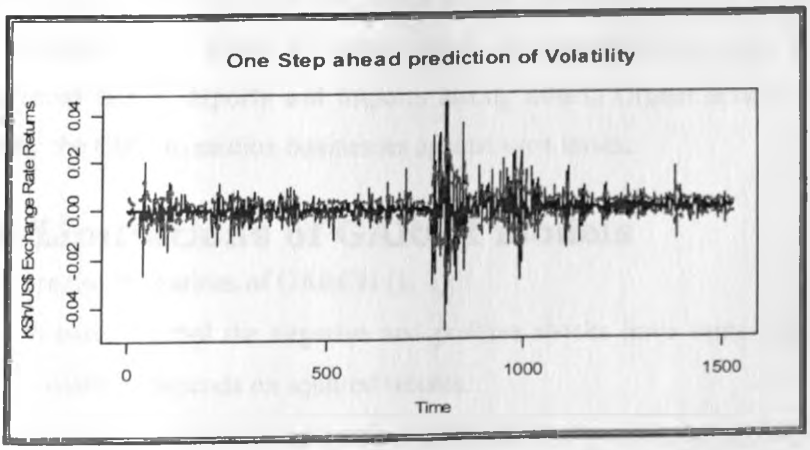
i.e. 382 days, where n is the number of observations in the dataset and x is a constant 0.25).

Figure 12: Predicting using GARCH (1,1) at 95% confidence



In figure 13 below one day a head prediction for GARCH with $\pm\hat{\sigma}_t$ is shown. Indeed the model created, from observations in this figure mimics the returns. Hence the model can be used to give an approximation of how the returns predictions at a time in the future. The GARCH model, in most cases of high volatility tends to under-predict the volatilities for the next day. In low volatility periods, the predictions seem to have a good prediction. Generally on average the model predicts the returns well.

Figure 13: One step prediction of Volatility



4.6 Conclusions

Exploratory analysis showed that the exchange rates are platykurtic and slightly negatively skewed. ARIMA(4,1,2) model was fitted on the exchange rate returns data by picking the model with the least AIC, AICC and BIC statistics. Then diagnostic checks done showed the residuals were white noise but exhibited volatility clustering which is an indication that the data is not homoscedacitic. GARCH(1,1) was then applied to counter the volatility clustering characteristic exhibited in the residuals of ARIMA(4,1,2) on the KES vs. US$ daily exchange returns for the period 2nd November 2004 to 31st December 2010. GARCH (1,1) fitted well on the data and indeed the residuals after the GARCH(1,1) fitting showed homoscedasticity and white noise properties

The Quasi-likelihood procedure used has parametric estimators that are consistent and asymptotically normal. The estimated models fit the data well, thereby confirming the empirical evidence in Bollerslev et al. (1992), that the GARCH (1,1) is adequate in describing volatility in many financial time series. Therefore, the knowledge of volatility and its estimation can ensure mitigation of the long term risk of any investment, (Choy, 2002). This in turn assists in promoting economic growth, since investment is the main channel of increasing real output and employment.

4.7 Recommendations

Prediction generated match the volatility that is being exhibited of late on the KES depreciation against the US dollar, which causes inflationary effects in the economy and generate great losses to many highly exposed business such as tourism, commercial banks, exports and imports among others. Urgent actions need to be taken by the CBK to caution businesses against such losses.

4.8 Limitations of GARCH models

Below are the limitations of GARCH (1, 1);

i. It assumes that the negative and positive shocks have same effect because volatility depends on squared returns.

ii. The likelihood is flat unless the number of observations is very large;

iii. The model tends to overpredict volatility because it responds slowly to large isolated returns.

Future researchers interested in the GARCH analysis can apply the other methods to address these limitations.

REFERENCES:

1. Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle in B. N. Petroc and F. Kaski, editors... Second International Symposium in Information Theory (pp. 267-281), Budapest: Akademiai Kiado.

2. Allayannis, G. I. (2001). Exchange Rate hedging: Financial vs. Operational Strategies. American Economic Review Papers and Proceedings, 91(2), (pp. 391-395).

3. Ardia, D. (2008). Financial Risk Management with Bayesian estimation of GARCH models. in D. Ardia, Theory and Applications. Berlin: Springer-Verlag.

4. Bollerslev, T., Chou, R.Y. and Kroner, K.F. 1992. ARCH modeling in finance. A review of the theory and empirical evidence. Journal of Econometrics 52, 5-59.

5. Central Bank of Kenya (2006). Annual report for 2006. Nairobi: Central Bank of Kenya.

6. Central Bank of Kenya (2008). URL: http://www.centralbank.go.ke/FAQs/MonetaryPolicyFAQs.aspx., accessed 17th July 2011.

7. Chan, J. C. (2008). Time Series Analysis, with Applications in R, 2nd edition. New York: Springer.

8. Chen (1997). Forecasting the S&P 500 index volatility International Review of Economics and Finance 6(4) pp.391-404.

9. Christian Francq, J.-M. Z. (2010). Garch Models: Structure, Statistical Inference and Financial Applications. West Sussex: Wiley.

10. Engle. R. (2001). GARCH 101: The use of ARCH/GARCH models in Applied Econometrics. Journal of Economic Perspectives Vol 15, Number 4 . 157-168.

11. H. Akaike. Information theory and an extension of the maximum likelihood principle. In B. N. Petroc and F. Kaski, editors. Second International Symposium in Information Theory, pp. 267-281, Akademiai Kiado, Budapest, 1973.

12. Isaya Maana, P. M. (2010). Modelling the volatility of exchange rates in the Kenyan Market. African Journal of Business Management. Vol 4(7), 1401-1408.

13. Jurgen Franke, e. a. (2008). Statistics of Financial Markets: An introduction. Berlin: Springer-Verlag.

14. Lessep M. and Morrel J. G., (1997). Forecasting Exchange Rate: Theory and Practice, The Henley Centre for Forecasting, London.

15. Lim. C and McAleer M (2002). Time series forecasts of international travel demand for Australia, Tourism Management, 23(4) pp 389-396.

16. Madura J , Martin A.D and Wiley M (1999) Forecast bias and accuracy of exchange rates in emerging markets, Journal of Multinational Financial Management 9 pp. 27-43.

17. Pagan A. and Schwert G. W., (1990). "Alternative models for conditional stock volatility", Journal of Econometrics, 45, 267-290.

18. R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna Austria. URL: http://www.R-project.org, accessed 20th July 2011.

19. Scott, K. J. (2000). Garch Models and the Stochastic Process Underlying Exchange Rate Price Changes. Journal of Financial and Strategic Decisions Vol 13 Number 2.

20. Vox (2011). Research-based policy analysis and commentary from leading economists, London, UK. URL: www.voxeu.org/index.php?q=node/4819, accessed 21st June 2011.

21. Xuezhheng Bai, J. R. (27 July 2001). Kurtosis of GARCH and Stochastic Volatility Models with Non-normal Innovations. Chicago.

APPENDICES

Appendix 1: Sample data of the KES/US$ exchange rate

Date	Average Exchange rate
11/2/2004	81.2078
11/3/2004	81.2117
11/4/2004	81.1472
11/5/2004	81.1306
11/8/2004	81.1444
11/9/2004	81.1500
11/10/2004	81.1439
11/11/2004	81.0689
11/12/2004	81.1139
11/16/2004	81.1139
11/17/2004	81.2111
11/18/2004	81.2411
11/19/2004	81.2944
11/22/2004	81.3389
11/23/2004	81.3500
11/24/2004	81.2400
11/25/2004	81.2444
11/26/2004	81.2278
11/29/2004	81.2278
11/30/2004	81.2422
12/1/2004	81.2506
12/2/2004	81.2461
12/3/2004	81.2494
12/6/2004	81.2278
12/7/2004	81.2333
12/8/2004	81.2267
12/9/2004	81.2294
12/10/2004	81.2011
12/14/2004	81.2278
12/15/2004	80.3611
12/16/2004	79.8744
12/17/2004	79.6356
12/20/2004	79.6578
12/21/2004	79.4889

Appendix 2: Program Code

Appendix 2.1: Exploratory Data Analysis

The R Statistical Packages used are fracdiff, tseries, moments, forecast, stats, fSeries, graphics, MASS, FinTS, FArma, TSA eGarch and fGarch (R Development Core Team (2011).

```
##importing the data from .csv file extension
prices=read.csv("exchange1.csv", header=T)
##Exploratory data analysis, the rates are in 2nd column..Exchange rate time plot fig
1
plot(prices[,2],type="l",xlab="days", ylab="KES/US$ rates", main="Time Series
KES/US$ rates")
##QQ plots and other normality analysis
qqnorm(prices[,2]);qqline(prices[,2])
shapiro.test(prices[,2])
summary(prices[,2])
kurtosis(prices[,2])
skewness(prices[,2])
###checking for stationarity of exchange rate prices
##acf and pacf of the rates fig 3
par(mfrow=c(1,2))
acf((prices[,2]),50)
pacf((prices[,2]),50)
##Augmented Dickey-Fuller test
adf.test(prices[,2])
###The returns
summary(diff(log(prices[,2])))
var(diff(log(prices[,2])))
kurtosis(diff(log(prices[,2])))
skewness(diff(log(prices[,2])))
```

Appendix 2.2 : Log- Differencing

```
##Model parameter estimation by log-differencing to gain stationarity
acf(diff(log(prices[,2])),50)
pacf(diff(log(prices[,2])),50)
#ADF test for returns
adf.test(diff(log(prices[,2])))
##test for normality
shapiro.test(diff(log(prices[,2])))
qqnorm(diff(log(prices[,2]))); qqline(diff(log(prices[,2])))
##the plot fig 4
plot(diff(log(prices[,2])), type="l",xlab="days", main="Returns")
win.graph(width=4.875, height=3,pointsize=8)
auto0=auto.arima(diff(log(prices[,2])))
summary(auto0)
```

```
#diagnistic tests for residuals pg 6
Box.test(auto0$residuals, type="Ljung")
par(mfrow=c(3,1))
plot(rstandard(auto0), main="Standardised residuals", type="l")
acf(auto0$residuals,100, main="ACF Standardised Residuals")
pacf(auto0$residuals,100, main="PACF Standardised Residuals")
##checking for characteristics that allude high-order dependency fig 7
returns=diff(log(prices[,2]))
par(mfrow=c(3,1))
acf(returns,30)
acf(returns^2,30)
acf(abs(returns),30)
Box.test(returns, type="Ljung")
Box.test(returns^2, type="Ljung")
Box.test(abs(returns), type="Ljung")
shapiro.test(auto0$residuals)
```

Appendix 2.3: Conditional Heteroscedasticity

```
##Testing for ARCH effects, fig 8
McLeod.Li.test(auto0,y=diff(log(prices[,2])))
###fitting Garch(1,1)
returns=diff(log(prices[,2]))
m1=garch(x=returns,order=c(1,1))
summary(m1)
##determining the properties of the GARCH fit, fig 9
plot(residuals(m1),type="l",ylab="standardized residuals", xlab="days")
qqnorm(residuals(m1));qqline(residuals(m1))
shapiro.test(residuals(m1))
plot(residuals(m1),ylab="standardized residuals")
#garch residuals autocorrelations
par(mfrow=c(3,1))
acf(residuals(m1),100,na.action=na.omit)
acf(abs(residuals(m1)),100,na.action=na.omit)
acf(residuals(m1)^2,100,na.action=na.omit)
Box.test(residuals(m1)^2,type="Ljung")
#checking for goodness of fit, fig 10 and 11
gBox(m1,x=returns,method='squared')
gBox(m1,x=returns,method='absolute')
#prediction, fig 12
fit0 = garchFit(~ garch(1, 1), data = returns, trace = F)
predict(fit0, n.ahead = 10, plot=TRUE, crit_val=1.96)
#One-step ahead prediction of volatility, fig 13
plot(1:1530,returns[1:1530],type="l",xlab="Time",ylab="KES/US$ Exchange Rate
Returns")
lines(fit0_predict[,1],col="red",lty="dashed",lwd=2)
lines(fit0_predict[,2],col="red",lty="dashed",lwd=2)
title(main="One Step ahead prediction of Volatility")
```