

**MODELLING RATES OF INFLATION IN KENYA:AN APPLICATION OF
GARCH AND EGARCH MODELS**

SAMMY OKETCH FWAGA

REG NO:SC384-C012-1458/2014

A Research Project Submitted In Fulfillment Of The Requirement For The Award Of The
Degree Of Master Of Science In Applied Statistics Of Jomo Kenyatta University Of Science
And Technology

DECLARATION

This thesis is my original work and has not been presented for a degree in any other University

Signature..... Date.....

Sammy Oketch Fwaga

This thesis has been submitted for examination with our approval as University
supervisors

Signature..... Dates.....

Prof. George Orwa

JKUAT, Kenya

Signature..... Date.....

Mr.Henry Athiany

JKUAT, Kenya

Signature..... Date.....

DEDICATION

To my wife Esther and daughter Naymah

ACKNOWLEDGMENT

This work has benefited from the selfless support that I got from many people. Suffice is to say that it would not have been possible to succeed without the support from my family. Most significantly my heartfelt gratitude goes to Mr. Henry Athiany and Prof. George Orwa my supervisors. Their constructive criticism made complex issues much simpler and clearer. Finally, I thank the Almighty God for the strength and courage to wade through the rigorous process while undertaking the study.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGMENT	iii
LIST OF FIGURES	vii
LIST OF TABLES	ix
DEFINITION OF KEY TERMS	x
ABSTRACT	xi
1 INTRODUCTION	1
1.1 Background of the study	1
1.2 Statement of the problem	5
1.3 Objectives of the study	6
1.3.1 General objectives	6

1.3.2	Specific objectives	7
1.4	Justification of the study	7
1.5	Limitations	8
1.6	Organization of the project	8
2	LITERATURE REVIEW	9
2.1	Introduction	9
2.2	Optimal model for predicting inflation data	10
2.3	Forecasting inflation data	13
3	METHODOLOGY	18
3.1	Model Selection Using Information Criteria	19
3.2	Statistical methods and model review	23
3.2.1	ARCH(q) Model	23
3.2.2	GARCH(1,1) Model	27
3.2.2.1	Properties of GARCH Model	28
3.2.3	EGARCH(p,q)	29
3.2.3.1	Forecasting with EGARCH model	30

3.2.3.2	Properties of EGARCH model	32
3.3	Parameter Estimation	32
4	EMPIRICAL STUDY	34
4.1	Source of data	34
4.2	Analysis Methods	35
4.3	Empirical Study	37
4.3.1	Testing for Serial Correlation and ARCH effects in the return series .	37
4.3.2	Model estimation and evaluation	40
4.3.2.1	Evaluation of estimated EGARCH (1, 1) Model	45
5	DISCUSSIONS, CONCLUSION AND RECOMMENDATION	52
5.1	Discussion	52
5.2	Conclusion	53
5.3	Recommendation	54

List of Figures

4.1	Graph of inflation return series	38
4.2	Partial autocorrelation function of return series	41
4.3	QQ Plot for normality	48
4.4	ACF of Squared Standard Residuals	49
4.5	Plot of residual against time	50

List of Tables

4.1	Data description of inflation and return series	36
4.2	Engle ARCH test for heteroscedasticity for return series	39
4.3	Engle ARCH test for squared for return series	39
4.4	Ljung Box-Q-test for returns	39
4.5	Ljung Box-Q-test for squared returns	40
4.6	Comparison of Garch and Egarch Models	42
4.7	Parameter Estimates for EGARCH Model	42
4.8	Parameter Estimates for GARCH(1,1) Model	43
4.9	Parameter Estimates for Garch(1,2) Model	44
4.10	Parameter Estimates for Garch(1,2) Model	44
4.11	Parameter Estimates for Garch(2,2) Model	45
4.12	Weighted Ljung-Box Test on Standardized Squared Residuals	47
4.13	Weighted ARCH LM Tests	47

4.14 Forecast Inflation Data	51
--	----

DEFINITION OF KEY TERMS

Volatility

This is the degree of variation of inflation series over time measured by the standard deviation of returns

Time Series

It is a series of inflation data recorded over time. In this case monthly inflation data recorded from January 1985 to December 2015

Inflation

This is the persistent increase in the general price level

Heteroscedasticity

This is the time varying variance of the inflation data under study

ABSTRACT

The effects of inflation on the economy are diverse and can both be beneficial or detrimental to the economy. The negative effects are however most pronounced and comprises a decrease in real value of money as well as other monetary variables. As a result uncertainty over future inflation rate may discourage consumers. It may also lead into decrease in foreign investments in a country. The purpose of this study was to determine an effective Arch-type model for forecasting Kenya's inflation. Using Kenya monthly inflation data from January 1990 to December 2015, the performance of GARCH and EGARCH type models were analyzed to come up with the best model for forecasting Kenyan inflation data. Since the inflation series is non-stationary, the Consumer Price Index (CPI) was first transformed to return series by logarithmic transformation. Afterwards, the data was tested for the presence of ARCH effects and serial correlation using both Ljung Box Pierce Q test and Engle Arch test. The test showed presence of heteroscedasticity and correlation in the inflation return series which is a key feature of a financial time series data. The project adopted AIC and BIC approach in selecting the the best model. From the fitted models EGARCH(1,1) had the smallest AIC and BIC values followed by the GARCH(1,1) model. Model diagnostic test was conducted on the selected model EGARCH (1,1) to determine its adequacy and goodness of fit. QQ plot was fitted to the residuals of the model and fairly straight line was produced looking roughly linear. Furthermore weighted Ljung Box Test on standard squared residuals showed the absence of correlation in the model. In conclusion, EGARCH(1,1) model is the best

model for forecasting Kenyan inflation data.

Chapter 1

INTRODUCTION

1.1 Background of the study

A great deal of data in business, economics, engineering and natural sciences occur in the form of the time series where observations close together in time domain are more correlated than observations further apart. Chatfield(2000) defines time series as a series or sequence x_t of data points measured typically at successive times. This study explored GARCH and EGARCH models to understand the inflation process and consequently the best model for forecasting inflation. Time series models are useful for forecasting and trend analysis in business and finance. In order to make future decisions regarding a given set of data, forecast is necessary. Major government and firms decisions are based on past historical data. Accordingly, time series involves explaining past observations in order to try and predict those in the future (Ahiati,2007). In the same manner Chatfield (2000) defines time series as measurements made continuously through time or taken at a discrete set of time points.

A significant macroeconomic issue that has posed a great challenge to most Nations and

states monetary authorities today is tracking and predicting reliably the movement in the general price level. Kenya like most developing countries has had significant challenges between policy formulation, policy implementation and policy targets. In most cases it is difficult to attain the targets due to insufficient models that can be used to forecast inflation and its real determinants.

Inflation is a major monetary policy performance indicator and is useful in informing the investors, the general public and government about the trends in movement of the currency. The investor would use the forecasted inflation rates to make better investments decisions. The knowledge of these indicators drives inflationary expectation and therefore serves as a nominal anchor for bargaining process and fixed contracts (Moser et al., 2004). As a result, a clear understanding of inflation forecasting technique is crucial for the success of monetary policy in tracking the movement of macroeconomic aggregates and in maintaining stable and sustainable economic growth.

When the general level of price is relatively stable the uncertainties of time-related activities such as investment diminish. The consequence is the promotion of full employment and robust economic growth. According to Sobel et al., (2006), when price stability is achieved and maintained, monetary policy makers have done their job well. Understandably government top priority is ensuring a healthy economy which promotes the well-being of her citizens. Government through its ability to tax, spend and control money supply, attempts to promote full employment, price stability and economic growth (Ezekiel et al., 2015).

Time series models have many forms, and represents different stochastic process which could

be linear and non-linear. Among the linear models include autoregressive (AR) model of order (p), moving average (MA) of order(q) and autoregressive moving average (ARMA) model of order (p,q). A combination of the above model produces the autoregressive moving average (ARIMA).

The non-linear time series model represent or reflect the changes of variance along with time known as heteroscedasticity. With these models, changes in variability are related to and/or predicted by recent past values of the observed series. The wide variety of non-linear models include the symmetric models such as Autoregressive Conditional Heteroscedastic (ARCH) model with order (p) and Generalized ARCH (GARCH) model with order(p,q). Other asymmetric models are the Power ARCH (PARCH), Threshold GARCH (TGARCH), Exponential GARCH (EGARCH), Integrated GARCH (IGARCH), etc. All these asymmetric models have order (p,q). The above mentioned non-linear models form part of a large family of the ARCH type models. In this study, two of such models , GARCH and EGARCH- would be fitted to the data set.

Autoregressive (AR), Moving Average (MA) and ARMA models are often very useful in modeling general time series. However, they all have the assumption of homoskedacity (or equal variance) for the errors. This may not be appropriate when dealing with financial markets variables such as stock price indices or inflation rate. Heteroscedacity affects the accuracy of forecast confidence limits and has to be handled properly by constructing appropriate non-constant variance models (Amos, 2009). Financial markets variables typically have the following three characteristics which general time series models have failed to consider: The distribution of a financial series X_t , has heavier tails than normal; values of X_t

do not have much correlation but values of X_t^2 are highly correlated. Finally the changes in X_t tend to cluster. Large (small) changes in X_t tends to be followed by large (small) changes, as documented by Mandelbrot(1963).

Inflation data can be studied using a stochastic modelling approach that captures time dependent structures embedded in the time series inflation data (Edward, 2011). The autoregressive conditional heteroscedasticity (ARCH) models, with its extension to the generalized autoregressive conditional heteroscedasticity. Numerous applications of the ARCH model have been reported. Eagle (1983) discussed the conditional variance of inflation rate using the ARCH model and found that the inflation tended to change over time. In Weiss (1984), ARMA model and the ARCH model were found to be incorporated successfully in modelling sixteen U.S macroeconomic time series. Lastrapes(1989) confirmed that the ARCH model can account for many of the empirical regularities of weekly exchange rate data. Bera and Higgins (1993) provided a discussion of major contribution to the ARCH model. They noted that the ARCH model was useful to capture various stylized properties of time series data such as the leverage effect in volatility, volatility clustering, excess kurtosis and heavy tails. However, the ARCH model has some weaknesses for higher order models. It turns out to be difficult to estimate parameters because the process requires long-lag length and a large number of parameters to be estimated (Zakaria and Abdalla, 2012). Another weakness of the ARCH model is that the process often produces negative coefficient estimators in parameters α_i . The more parameters to be estimated in the ARCH model, the more likely it can obtain a negative estimated value.

In order to avoid the long lag structure of the ARCH model and solve negative coefficient

problem, a Generalized ARCH (GARCH) model was developed by Bollersleve (1986). The GARCH model has been modified to accommodate the possibility of serial correlation in volatility. It contains a linear combination of lags of the squared residuals from the conditional variance (Goudarzi, 2010).

Empirical studies have established that the GARCH is a more parsimonious model than the ARCH model (Poon and Granger, 2003) and becomes a valuable model for volatility forecasting in financial time series. In particular the GARCH (1, 1) is the most popular model for estimating and forecasting volatility. The ARCH model incorporates the autoregressive term in return series whereas the GARCH model is superior to ARCH because it adds the general feature of conditional heteroscedasticity terms. The conventional GARCH model, which is called symmetric GARCH model is not always a perfect model and could be improved because the error terms are assumed to be normally distributed. Consequently, the symmetric GARCH model is less than adequate to fully account for some stylized fact of returns. It is as a result of such limitation that the study will also incorporate EGARCH model in order to improve volatility forecasting in cases of financial stylized characteristics of return series.

1.2 Statement of the problem

The assumption of constant variance over some period when the resulting series moves or progress through time, statistically is inefficient and inconsistent (Campbell, et al., 1997). While dealing with financial data for instance, inflation data variances changes with time a

phenomenon defined as heteroscedasticity, hence there is a need of studying models which accommodates possible variations in variance.

Considering the issues of inflation modelling and forecasting in Kenya, no previous statistical or mathematical study has used Kenyan inflation rate data to model the conditional variance and forecast future inflation values. As a result, this study intends to use autoregressive conditional heteroscedasticity (ARCH) models with its extension to the generalized ARCH (GARCH) models to model the monthly inflation data and determine if its best suited for forecasting financial data based on past observation. Taking into consideration the limitation of conventional GARCH model since it is only symmetric and cannot explain asymmetric variation in inflation data EGARCH model will also be used as to improve volatility forecasting and hence determining the best model.

1.3 Objectives of the study

1.3.1 General objectives

The main objective of the study is to establish the best model for forecasting inflation data between the two selected models of GARCH and EGARCH.

1.3.2 Specific objectives

1. Fit each of the two times series model (GARCH and EGARCH) using Kenya inflation rates data
2. Identify the optimal model
3. Predict a one year out-sample forecast based on the optimal model

1.4 Justification of the study

The research study is of great significance in both academic research, economic development and business as follows: In terms of economic development, the government will be in a position to design suitable fiscal and monetary policies when they are in a position to predict with reasonable accuracy future inflation rates. This will be vital to the central bank in creating a conducive business environment and thus enhancing a country's balance of payment and balance of trade. The accuracy of the forecasting will also allow the central bank to control liquidity in all the commercial banks by establishing suitable bank reserve. As far as business is concerned, investors will have confidence in a country when there is stable economic climate and certainty in a country's inflation rate. Finally in terms of research, the outcomes will be useful in adding to the body of knowledge as well suggesting areas that require further research.

1.5 Limitations

This research work is limited to establishing a suitable model for forecasting Kenyan inflation data, however it does not go beyond the key factors as reported in the dataset. The key factors derived herein therefore pertains to the factors of this very data set. Moreover, GARCH models are parametric specifications that operate best under relatively stable market conditions. Although GARCH is explicitly designed to model time-varying conditional variances, GARCH models often fail to capture highly irregular phenomena; including wild market fluctuations.

1.6 Organization of the project

The rest of this project is organized as follows: In the subsequent chapters, there is literature review in which literature relating to the specific objectives have been reviewed and gaps therein exposed through a critique. Thereafter, the chapter that follows has the methodology in which the statistical method and model have been discussed. The following chapter after methodology has the empirical study in which the actual data analysis has been carried out. Finally there is findings conclusions and recommendations.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

This chapter look into the relevant theories and concepts associated with inflation and related studies that has been carried out by other researchers on the topic area. The chapter is divided into two main sections namely: The Concept of inflation and Review of related works.

Concept of Inflation

According Webster (2000), inflation is the persistent and continuous rise in the levels of the consumer prices in an economy. Inflation can also be seen as the persistent decline in the purchasing power of money. That is, inflation means that your money cannot buy today as much as what it could have bought yesterday. Accordingly, different theories have been proposed by economists to explain the concept of inflation. These numerous theories can be grouped into two main broad theories; the excess- demand theory and the cost- push theory. The excess-demand theory argued that the excess demand for goods and services over supply in the economy is the main source of inflation as expressed by Hall (1982). On

the other hand, the cost-push theory of inflation asserts that inflation can be caused by the increase in the cost of production of firms. The increase in the cost of production will affect the profit margins of these firms and hence they will have to pass on the extra to consumers by increasing the prices of their products. The effects of inflation include among other things, people losing confidence in the currency as the real value of the currency is severely reduced. Also investors are scared to invest in a country due to the cost of doing business. Furthermore inflation can also lead to the wage-price spiral. This is the situation in which there are higher wage demands as people try to maintain their real living standards. This leads businesses to increase prices to maintain profits and higher prices then put further pressure on wage.

Bailey (1956) observed that inflation has negative effects on the economy through its cost on welfare. He also stated that the costs associated with unanticipated inflation are the distributive effects from creditors to debtors, increasing uncertainty affecting consumption, savings, and borrowing and investment decisions. From the foregoing statements it is clear that coming up with a suitable model to forecast future inflation is key for purposes of planning and consequently mitigating the negative impact as a a consequence.

2.2 Optimal model for predicting inflation data

Perniagaan, Ekonomi and Nor (2004) explored the varying volatility dynamics of inflation rate data in Malaysia for the period from August 1980 to December 2004. The Generalized autoregressive conditional heteroscedasticity (GARCH) models and the exponential general-

ized autoregressive conditional heteroscedasticity (EGARCH) models were used to capture the stochastic variations and asymmetries in the data. An in-sample evaluation of the sub-periods volatility was done. Based from the results using both models the GARCH model of order (1, 1) and EGARCH model of order (1, 1) produced good estimates of sub-periods volatility. But due to nature of the data of having highly irregular fluctuation EGARCH model was selected to be the best.

According to Garcia, et al (2003), General Autoregressive Conditional Heteroskedasticity (GARCH) models consider that the price series is not invariant (i.e. the error term: real value minus forecasted value does not have 0 mean and constant variance). The error term is now assumed to be serially correlated and can be modeled by an Autoregressive (AR) process. Thus, a GARCH process can measure the implied volatility of a series due to price spikes. For example, California experienced huge price spikes during the summer of 2000 that led to the closure of the market until new rules were developed. As suggested by Bollerslev et al. (1994), economic loss functions that explicitly incorporate the costs faced by volatility forecast users provide the most meaningful forecast evaluations.

Moreover Vosvrda and Zikes (2004) studied the behaviour of volatility and the distributional properties of the Czech, Hungarian and Polish stock markets data for the period 1996-2002 period using weekly data, they use PX-50 index for Czech republic and found statistically significant results for GARCH (1,1) model and concluded that the volatility of the returns on PX -50 is very persistent. According to Rockinger and Urga (2012) while investigating two groups of countries which are transition economies and established economies. Transition economies are; Czech Republic, Poland, Hungary and Russia, established economies are

USA, Germany and UK. The model results are very similar for Czech Republic, Hungary, and Poland for all countries investigated there exists significant Garch effect.

According to Awartoni and Carradi (2005) who compared the relative predictive ability of the GARCH (1,1) model with the alternative asymmetric GARCH models, such as EGARCH, GJR-GARCH, TGARCH and Asymmetric GARCH, AGARCH model. They used the daily Standard and Poor 500 composite price index to model and forecast volatility at 1-,5-,10-,15-,20-,and 30 steps ahead. All models were considered under the normal error distribution. They found that Asymmetric GARCH models were superior to GARCH (1, 1) models for all prediction horizon. Further, Su (2010) employed both GARCH and EGARCH models in studying the financial volatility in China. He applied the daily stock returns data from January 2000 to April 2010 and split the time series into two parts: before the crisis and during the crisis period. The empirical results suggested that EGARCH model fits the sample data better than GARCH model in modelling the volatility of Chinese stock returns. The result also showed that long term volatility was more volatile during the crisis period whilst bad news produced stronger effect than good news for the Chinese stock market during the crisis.

Ngailo (2011) modelled monthly inflation data for Tanzania form the year 1997 to 2010 using GARCH model. He found out that GARCH (1, 1) model was the best model for forecasting the Tanzania inflation data. However in his study, one of the assumption he made was that the inflation data was symmetric and therefore he did not attempt to employ other ARCH type models that would cater for any asymmetry in the data in case it existed. This study will try to explore both ARCH-type models that can accommodate asymmetry and

symmetry pattern that might be exhibited from the Kenyan inflation data.

2.3 Forecasting inflation data

Albert et al. (2006) investigated the forecasting performance of GARCH (1, 1), EGARCH (1, 1), GJR-GARCH (1, 1) and Asymmetric power ARCH (APARCH(1,1)) models with different error distributions: normal, student-t, and skewed student t. The result showed that the EGARCH model using a skewed student-t distribution was the most successful in volatility forecasting. Shamiri and Isa (2009) examined the relative efficiency of three different types of GARCH models in terms of their volatility forecasting performance. They compared the performance of symmetric GARCH (1, 1), asymmetric EGARCH (1, 1) and non-linear asymmetric NAGARCH (1, 1) models with six error distributions namely normal, skewed normal, student-t, normal inverse Gaussian and generalized error distribution. They found out that EGARCH model provided better performance on volatility forecasting than the GARCH (1, 1) model.

Jean-Philippe (2001) examined the forecasting performance of four GARCH-typed models. The comparison focused on two different aspects; the difference between symmetric GARCH model (traditional GARCH model) and asymmetric models (EGARCH, GJR and APARCH) and the difference between normal tailed symmetric, fat-tailed symmetric and fat tailed asymmetric distributions (i.e. normal distributions against student-t and skewed student-t distributions). The study concluded that noticeable improvements were made when using an asymmetric GARCH in the conditional variance and that the APARCH and

GJR outperformed the EGARCH. Furthermore, non-normal distributions provided better in-sample results than Gaussian distributions.

According to Brooks (2008) in his study of stochastic volatility models found that most time series models such as GARCH, will have forecasts that tend towards the unconditional variance of the series as the prediction horizon increases. This is a good property for a volatility forecasting model to have, since it is well known that volatility series are mean-reverting. This implies that if they are at a low level relative to their historic average they will have a tendency to rise back towards the average. This feature is accounted for in GARCH volatility forecasting models.

Yuksel and Bayram (2005) investigated the stock market volatility in Turkish, Greek and Russian stock markets using the total return indexes based on the domestic currencies of the corresponding countries. The data set covers a period from 1994 -2004. The study concluded that the GARCH-M (1,1) was the best model for modelling the volatility in the stock markets in Turkey. In the case of the stock markets of Greece, the TARARCH (1,2) was the best model whilst the TARARCH (1,1) was the best model for the Russian stock markets. Anna (2011) examined the relationship between inflation, inflation uncertainty and output growth with evidence from the G-20 countries using several GARCH and GARCH-M models in order to generate a measure of inflation uncertainty. The study adopted two approaches to test for the impact of inflation uncertainty on inflation and vice versa. The first approach was based on the GARCH-M model that allows for simultaneous feedback between the conditional mean and variance of inflation while the second approach was based on a two-step procedure where Granger methods were employed using the conditional variance of a

simple GARCH model. The results of the study suggested significant positive relationship between inflation uncertainties and inflation in most countries. These results go to support the Cukierman-Matter and Friedman-Ball hypothesis. Also the results of the study provided evidence for the Holland theory; that uncertainty lead to lower and in the case of the effect of inflation uncertainty in output growth, there was little evidence that inflation uncertainty has negative real effects.

Chatfield (2000) asserted that the idea behind a GARCH model was similar to that behind the ARMA model with respect to the fact that a higher order AR or MA model may often be approximated by a mixed ARMA model with fewer parameters using a rational polynomial approximation. He described the GARCH model as an approximation to a higher-order ARCH model. He noted that the GARCH (1, 1) model has become the standard model for describing non constant variance due to its relative simplicity. Empirical evidence has revealed that often $(\alpha+\beta)\leq 1$ so that the stationarity condition may be met. However if the $(\alpha +\beta) = 1$, the process ceases to have a finite variance although it can be shown that the squared observations are stationary after taking first differences. In such a situation a better model Integrated GARCH (IGARCH) developed by Engle and Bollerslev (1986) is recommended.

Hajek (2007) tests the Efficient market Hypothesis on the Czech Capital market for 1995-2005 period for monthly, weekly and daily data. The author analyses efficiency and linear dependency of several index closing values and stock closing prices on the Prague Stock Exchange and concludes that both daily stock returns and daily index returns are significantly linearly dependent and the hetereskedasticity-consistent methodology must be therefore ap-

plied to avoid significant biases.

In his study of three widely used volatility models Minkah (2007) examined the forecasting ability of the three models namely the Historical Variance, the Generalized Autoregressive Conditional Heteroscedastic (GARCH) Model and the Risk Metrics Exponential Weighted Moving Average (EWMA). The characteristics of these volatility models were explored using data on the Standard and Poors 500 Index, Dow Jones Industrial Average (DJIA), OMX Swedish Stock Exchange (OMXS30) index, Dow Jones-AIG Commodity Index (DJ-AIGCI), the 3 Months US Treasury Bill Yield, the Ghanaian Cedi and the US Dollar (CEDI/USD) exchange rates.

It was observed that the complex models i.e. GARCH (1, 1) and Risk Metrics EWMA outperformed the simple Historical Variance in the In-Sample volatility forecasts. The Out of Sample forecasting accuracy comparisons also revealed that for shorter forecasting horizons, the GARCH (1, 1) performed better whereas at longer horizons the simple Historical Variance outperformed all in most markets. This was due to the fact that complex models have more parameters and thus add to the estimation errors and its forecasts are consistently poor in Out-of-Sample.

Amos (2009) studied financial time series modeling using inflation data spanning from January 1994 to December 2008 for South Africa. In the study two time series models which are the seasonal autoregressive integrated moving average (SARIMA) model and the generalized autoregressive conditional heteroscedasticity (GARCH) model were fitted to the data for encountering trend and seasonal terms and accommodating time varying variance respec-

tively. A best fitting model for each family of models offering an optimal balance between goodness of fit and parsimony was selected. The SARIMA model of order (1, 1, 0) (0, 1, 1) and GARCH Model of order (1, 1) were chosen to be the best fitting models for determining the two years forecasts of inflation rate of South Africa. However GARCH model of order (1, 1) was observed to be superior in producing future forecasts because of its ability to capture variations in the data.

Engle, (1982) used the ARCH processes to model inflation rates recognizing that the uncertainty inherent in inflation tends to change with time. Talke (2003) in his MSc research also applied ARCH models to modeling volatility in time series data. In his work Talke, (2003) applied the ARCH methods to changes in exchange rates of the South African rand against the US dollar, Swiss and the British pound.

Chapter 3

METHODOLOGY

In determining the best model for predicting Kenyan inflation data, GARCH and EGARCH model was used. A rigorous model review review was carried out and subsequent model diagnostic to determine the optimal model determined.

Transformation

Most applied time series are non-stationary, their variance is changing with time. The inflation series for this study is non-stationary. Therefore the Consumer Price Index (CPI) was converted to returns series by logarithmic transformations. The formula for logarithmic transformation is:

$$R_t = \ln \frac{p_t}{p_{t-1}} \quad (3.0.1)$$

Where r_t is the return for any time t

p_t is the consumer price index value at time t

p_{t-1} is the consumer price index value at time t-1

Testing for ARCH effects and serial correlation in the return series

Ljung Box and Engle statistical tests were employed to find out the presence of serial correlation and ARCH effects in the data. Conducting various diagnostic tests is an important step in time series modelling. There exists numerous diagnostic tests designed to examine the dependence (correlation) structure of a time series. If a time series is serially uncorrelated, no linear function of the lagged variables can account for the behavior of the current variable. Ljung-Box-Pierce Q-test and Engles ARCH was used to test for heteroscedasticity. Conducting various diagnostic tests is an important step in time series modelling.

3.1 Model Selection Using Information Criteria

Model selection is a vital part of statistical forecasting. Accurate forecasting is only possible if one identifies an appropriate model. Model selection involves the use of information criteria to identify the best fitting model from a set of competing models. In this case the best fitting model is the model with the smallest value of the information criteria. The following are the two criteria used for selecting the best model:

Akaike Information Criteria(AIC)

This is one of the most commonly used criteria for model selection. It is a statistical measure of the likelihood of the set of competing models, penalized by the number of parameters in the model. AIC is defined as:

$$AIC = -2\log(L) + 2p \quad (3.1.1)$$

Where L is the likelihood under the fitted model. The smaller the AIC values, the better the model will be.

Bayesian Information Criteria

Another information criteria is called BIC. It is derived based on Bayes factors. BIC is defined as:

$$BIC = -2\log(L) + p\log(T) \quad (3.1.2)$$

The BIC criterion is different from AIC only in the last term. BIC depends on sample size T and p while AIC depends on p only. It is critical to decide which criterion will be appropriate to use for model selection. Markon and Krueger (2004) claimed that the ability for model selection of AIC performs well for small samples sizes, but does not perform well for larger sample sizes. BIC performs well with larger sample sizes while AIC seems to be more popular to use as a measurement for model selection. The advantages of the Bayesian information criterion is that for a wide range of statistical problems, it is order consistent (i.e. when the sample size goes to infinity, the probability of choosing the right model converges to unity) leading to more parsimonious model. Also, like the AIC, the BIC can be used to

compare in-sample or out-of-sample forecasting performance of a model. For the purpose of this project, both AIC and BIC will be used.

Model diagnostic checks and adequacy

The model diagnostic checks was conducted to determine the adequacy or goodness of fit of a selected model. The model diagnostic checks are performed on residuals and more specifically on the standardized residuals (Talke, 2003). The residuals are assumed to be independently and identically distributed following a normal distribution (Tsay, 2002). Plots of the residuals such as the histogram, the normal probability plot and the time plot of residuals can be used. If the model fits the data well the histogram of residuals should be approximately symmetric. The normal probability plot should be a straight line while the time plot should exhibit random variation (Bowerman and OConnell, 1997). The ACF and the PACF of the standardized residuals are used for checking the adequacy of the conditional variance model.

The Ljung Box Q-test (given are used to check the validity of the ARCH effects as well as test for autocorrelation in the data. To test the presence of ARCH effects, the null hypothesis of no ARCH effects is rejected if the significance probability value (p-value) is less than specified level of significance. In case of testing for the presence of autocorrelation, the null hypothesis of no autocorrelation is rejected if the Ljung Box (Q) statistics. Thus if the probability value of Ljung Box (Q) statistics of some of the lags are less than the specified level of significance, then the null hypothesis of no autocorrelation is rejected. Once the estimated model satisfies all these model assumptions, it can be seen as an appropriate representation of the data. Having established that the model fits the data well, the model can then be used to compute

forecasts of the series under consideration.

Assessment of predictiveness or forecast accuracy of the model

Forecast accuracy test can be used as criteria for selecting the best model. Several measures for assessing the forecast accuracy of ARCH-type models have been proposed. Some of these measures are the mean square error (MSE), mean absolute error (MAE) and Theils U statistic. The MSE is defined as the average of the squared difference between the actual variance and the volatility forecast ($\delta^2 t$). In the absence of the observed true variance, the squared time series observation $x^2 t$ is used. The MSE is given by:

$$MSE = \sum_{i=1}^T \frac{(x^2 t - \delta^2 t)^2}{T} \quad (3.1.3)$$

t=1: T

Where $\hat{\delta}_t^2$ t = 1 \dots T is the estimated conditional variance obtained from fitting the ARCH-type model. The MSE is criticized because x_t^2 is noisy and unstable and although the x_t^2 is a consistent estimator of δ_t^2 (Tsay,2002).Alternatively, other measures have been proposed.

The mean absolute error (MAE) was proposed by Lopez (1999) and defined as:

$$MAE = \sum_{t=1}^T \frac{|x_t^2 - \delta_t^2|}{T} \quad (3.1.4)$$

3.2 Statistical methods and model review

3.2.1 ARCH(q) Model

let r_t be the mean corrected return or rate of inflation, ϵ_t be the Gaussian white noise with zero mean and unit time t given by $H_t=(r_1, r_2, \dots, r_{t-1})$. Then the process (r_t) is ARCH(q) (eagle, 1982).

$$r_t = \delta_t \epsilon_t \quad (3.2.1)$$

$$\text{Where } E(r_t | H_t) = 0 \quad (3.2.2)$$

$$\text{Var}(r_t | H_t) = \delta_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i r_{t-i}^2 \quad (3.2.3)$$

$$\text{and the error term is such that } E(\epsilon_t | H_t) = 0 \quad (3.2.4)$$

$$\text{Var}(\epsilon_t | H_t) = 1 \quad (3.2.5)$$

Equations above show that the error term ϵ_t is a conditionally standardized martingale difference defined as follows: A stochastic series r_t is a martingale difference if its expectation with respect to past values of another stochastic series x_i is zero, that is $E(r_{t+i} | x_i, x_{i-1}, \dots) = 0$ for $i=1, 2, \dots$. In this type of the impact of the past on the present volatility it is assumed to be a quadratic function of lagged innovation.

Forecasting with the ARCH model

A very important use of the ARCH model is the evaluation of the accuracy of forecasts. In

standard time series methodology which uses conditionally homoskedastic ARMA process, the variance of the forecast error does not depend on the current information set. If the series forecasted displays ARCH, the current information set can indicate the accuracy by which the series can be forecasted. Eagle and Craft(1983) were the first to consider the effect of ARCH on forecasting. Baille and Bollerslev(1992) extended many of their results.

Let $r_1, r_2, r_3, \dots, r_t$ be an observed time series, then the 1-step ahead forecast, for $\ell=1, 2, \dots$, at the original t , denoted as $r_t(\ell)$, is taken to be the minimum mean squared error predictor, that is, $r_t(\ell)$ minimizes

$$E(r_{t+\ell} - f(r))^2$$

Where $f(r)$ is a function of the observation then

$$r_t(\ell) = E[r_{t+\ell} | r_1, \dots, r_t] \tag{3.2.6}$$

(Tsay, 2002)

However for the ARCH(1) model

$$r_t(\ell) = E[r_{t+\ell} | r_1, \dots, r_t] = 0 \tag{3.2.7}$$

The forecasts for the r_t series provide no much helpful information. It is therefore important

to consider the squared returns r_t^2 given as

$$r_t^2(\ell) = E[r_{t+\ell}^2 | r_1^2, \dots, r_t^2] \quad (3.2.8)$$

(Shephard,1996)

Hence the ℓ step ahead forecast for r_t^2 is given by

$$r_t^2(1) = \hat{\alpha}_0 + \hat{\alpha}_1 r_t^2 \quad (3.2.9)$$

which is equivalent to

$$\delta_t^2(1) = E[\delta_{t+1}^2 | r_t] \quad (3.2.10)$$

$$\hat{\alpha}_0 + \hat{\alpha}_1 r_t^2$$

Where α_0 and α_1 are the conditional maximum likelihood estimate of α_0 and α_1 . similarly a 2 step ahead forecast for r_t^2 is given by

$$r_t^2(2) = E[r_{t+2}^2 | r_t] \quad (3.2.11)$$

$$= E[\delta_{t+2} | r_t]$$

$$\begin{aligned}
&= \hat{\alpha}_0 + \hat{\alpha}_1 E[r_{t+1}^2 | r_t] \\
&= \hat{\alpha}_0 + \hat{\alpha}_1 (\hat{\alpha}_0 + \hat{\alpha}_1 r_t^2) \\
&= \hat{\alpha}_0 (1 + \hat{\alpha}_1) + (\hat{\alpha}_1^2 r_t^2) \\
&= \delta_t^2(2)
\end{aligned}$$

In general the ℓ -step ahead forecast from the r_t^2 is given by,

$$r_t^2(\ell) = E[r_{t+\ell}^2 | r_t] \tag{3.2.12}$$

$$= \hat{\alpha}_0 (1 + \hat{\alpha}_1 \hat{\alpha}_1^2 + \dots + \hat{\alpha}_1^{\ell-1}) + \tag{3.2.13}$$

$$= \delta_t^2(\ell) \tag{3.2.14}$$

The major weakness of using the ARCH model is that it can lead to a high parametric model if lag q is large **GARCH(p,q) Model** The generalized ARCH (GARCH) model was developed by Bollerslev(1986) as an extension of the ARCH model in the same way the ARMA process is an extension of the AR process. The principle of parsimony may be violated when a model has a large number of parameters resulting in difficulties in using the model to adequately

describe the data (Mbeah,2013). More importantly, the ARCH model may require many parameters as there might be a need for a large value of lag q. As a consequence, the principle of parsimony would be violated. A GARCH model is better than ARCH model when the principle of parsimony is applied.

generalized Autoregressive Conditional Heteroskedasticity (The GARCH(p,q))model of Bollerslev(1986) include plags of the conditional variance in the linear ARCH(q)

$$\delta_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_q r_{t-q}^2 + \beta_1 \delta_{t-1}^2 + \dots \beta_p \delta_{t-p}^2 \quad (3.2.15)$$

conditions on the parameters to ensure that the GARCH(p,q) contional variance is always positive are given in Nelson and Cao(1992). The GARCH(p,q) model may alternative be represented as an ARMA(max(p,q),p)model for the squared innovations.

3.2.2 GARCH(1,1) Model

GARCH(1,1) model is a particular case of GARCH(p,q) model where p and q are both equal to one. let r_t be the mean corrected return , ϵ_t , be a Gaussian white noise with mean zero and unit variance. Let also K_t be the information set or history at time t given by

$K_t=r_1, r_2, \dots, t_{t-1}$. The the process r_t is a GARCH(1,1) if

$$r_t = \delta_t \epsilon_t, \epsilon_t(0, 1) \quad (3.2.16)$$

$$\delta_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \delta_{t-1}^2 \quad (3.2.17)$$

Clearly from equations 4.1 and 4.2 a large past mean corrected squared returns r_{t-1}^2 or past conditional variance give rise to large values of δ_t^2 (Tsay,2002) where α_0, α_1 and β_1 are the parameters of the model such that $\alpha_0 \geq 0, \alpha_1 \geq 0, \beta_1 \geq 0$ and $(\alpha_i + \beta_i) \leq 1$ The constraints on the parameters are to ensure that the conditional variance δ_t^2 is positive as observed by (Mbeah,2013)

Forecasting with GARCH (p,q) Model

Forecasting of a GARCH model can be obtained using methods similar to those of an ARMA model. Thus the conditional variance of x_t is obtained by taking the conditional expectation of the squared mean corrected returns (Mbeah,2013). Consider the GARCH(p,q). Assuming a forecasting origin of time t, then the z-step ahead volatility forecast is given by:

$$x_t^2(z) = E(X_t^2 + z|x_t) \quad (3.2.18)$$

3.2.2.1 Properties of GARCH Model

The properties of GARCH model is similar to that of the ARCH model but requires far less parameters to adequately model the volatility process. The GARCH model is able to

model the volatility clustering but does not address the problem with the ARCH models lack of ability to model the asymmetric effect of positive and negative returns. The GARCH model also imposes restrictions on the parameters to have a finite fourth moment as is the case with ARCH model. The GARCH model is similar to the ARCH model but with the addition of lagged conditional variance as well as lagged squared returns. The addition of lagged conditional variances avoids the need for adding many lagged squared returns as is the case for the ARCH model to be able to appropriately model the volatility. This greatly reduces the number of parameters that need to be estimated.

3.2.3 EGARCH(p,q)

The exponential GARCH model was proposed by Nelson(1991)to overcome some weakness of the GARCH model in dealing with financial time series. The significant advantage of the EGARCH model is that if the parameters are negative, δ_t^2 will be positive. The nature logarithmic of the conditional variance is assumed as a linear function of its own lagged term and allowed to vary over time (Nelson,1991).

EGARCH(p,q) is given as:

$$\log \delta_t^2 = c + \sum_{i=1}^p g(Z_{t-i}) + \sum_{i=1}^q \beta_j \log \delta_{t-j}^2 \quad (3.2.19)$$

3.2.3.1 Forecasting with EGARCH model

The forecasting using an EGARCH model is done using methods similar to those of an ARMA model since the EGARCH model uses the usual ARMA parameterization to describe the evolution of the conditional variance of x_t . The simple EGARCH(1,1) model would be used to illustrate the multi-step ahead forecasts of the EGARCH models.

Assuming that the model parameters are known and the distributional assumption made on the error term is the standard Gaussian, the EGARCH (p,q) model is given as:

$$\ln \delta_t^2 = (1 - \alpha_1)\alpha_0 + \alpha_1 \ln(\delta_{t-1}^2) + g(\epsilon_{t-1}) \quad (3.2.20)$$

$$\text{where } g(\epsilon_{t-1}) = \theta_{\epsilon_{t-1}} + \gamma[|\epsilon_{t-1}| - E(|\epsilon_{t-1}|)] \quad (3.2.21)$$

$$= \theta_{\epsilon_{t-1}} + \gamma[|\epsilon_{t-1}| - \sqrt{\frac{2}{\pi}}] \quad (3.2.22)$$

Assuming a forecasting origin of t , then the 1-step ahead volatility forecast is given by

$$\delta_t^2(1) = \delta_{t+1}^2 = \delta_t^{2\alpha_1} \cdot \exp[(1 - \alpha_1)\alpha_0] \exp[g(\epsilon_t)] \quad (3.2.23)$$

Where all of the quantities on the right-hand side are known. For a 2-step ahead volatility

forecast , at the forecast origin t, equation(4.12) gives:

$$\delta_t^2(2) = \delta_{t+2}^2 = \delta_{t+1}^{2\alpha_1} \cdot \exp[(1 - \alpha_1)\alpha_0] \exp[g(\epsilon_{t+1})] \quad (3.2.24)$$

Taking conditional expectation on equation (4.13) at time t, we have

$$\delta_t^2(2) = \delta_{t+2}^2 = \delta_{t+1}^{2\alpha_1} \cdot \exp[(1 - \alpha_1)\alpha_0] E_t\{\exp[g(\epsilon_{t+1})]\} \quad (3.2.25)$$

Where E_t denotes a conditional expectation taken at the time origin t. The prior expectation can be obtained as follows:

$$E\{\exp[g(\epsilon)]\} = \int_{-\infty}^{\infty} \exp[\theta\epsilon + \gamma(|\epsilon| - \sqrt{\frac{2}{\pi}})] f(\epsilon) d\epsilon \quad (3.2.26)$$

$$= \exp(-\gamma\sqrt{\frac{2}{\pi}}) \left[\int_0^{\infty} e^{(\theta+Y)\epsilon} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{\epsilon^2}{2}} d\epsilon + \int_{-\infty}^0 e^{(\theta-Y)\epsilon} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{\epsilon^2}{2}} d\epsilon \right] \quad (3.2.27)$$

$$= \exp(-\gamma\sqrt{\frac{2}{\pi}}) \left[e^{\frac{(\theta+Y)^2}{2}} \Phi(\theta + Y) + e^{\frac{(\theta-Y)^2}{2}} \Phi(Y - \theta) \right] \quad (3.2.28)$$

Where $f(\epsilon)$ and $\Theta(x)$ are the probability density function and cumulative density function of

the standard normal distribution respectively. Consequently, the 2-step ahead volatility is

$$\delta_t^2(2) = \delta_{t+2}^2 = \delta_t^{2\alpha_1} \cdot \exp[(1 - \alpha_1)\alpha_0 - \sqrt{\frac{2}{\pi}}]x \left\{ \exp\left[\frac{(\theta + Y)^2}{2}\right] \Phi(\theta + Y) + \exp\left[\frac{(\theta - Y)^2}{2}\right] \Phi(Y - \Theta) \right\} \quad (3.2.29)$$

3.2.3.2 Properties of EGARCH model

The EGARCH model requires no restrictions on the parameters to assure that the conditional variance is non-negative. The EGARCH model is able to model volatility persistence, mean reversion as well as the asymmetrical effect. To allow for positive and negative shocks to have different impact on the volatility is the main advantage of the EGARCH model compared to the GARCH model.

3.3 Parameter Estimation

The log-likelihood function of the EGARCH(p,q) model is

$$L_t = c - \left(\frac{1}{2}\right) \sum_{t=1}^T \ln h_t - \left(\frac{1}{2}\right) \sum_{t=1}^T \left(\frac{\varepsilon_t^2}{h_t}\right) \quad (3.3.1)$$

$$\ln h_t = \alpha_0 + \sum_{j=1}^q \{ \alpha_j Z_{t-j} + \psi_j (|Z_{t-j}| - E|Z_t|) \} + \sum_{j=1}^p \beta_j \ln(h_{t-j}) \quad (3.3.2)$$

$$\text{Let } \beta = (\alpha_0, \alpha_1, \dots, \alpha_q, \psi_1, \dots, \psi_q, \beta_1, \dots, \beta_p) \quad (3.3.3)$$

The first partial derivative with respect to the EGARCH parameters are

$$\sum_{t=1}^T \frac{\partial L_t}{\partial \beta} = \left(\frac{1}{2} \sum_{t=1}^T \left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \right) \frac{\partial \ln h_t}{\partial \beta} \quad (3.3.4)$$

$$(3.3.5)$$

Where

$$\frac{\partial \ln h_t}{\partial \beta} = X\beta_t - \left(\frac{1}{2} \right) \sum_{j=1}^q \{ \alpha_j Z_{t-j} + \psi_j |Z_{t-j}| \} \frac{\partial \ln h_{t-j}}{\partial \beta} + \sum_{j=1}^p \beta_j \frac{\ln h_{t-j}}{\partial \beta} \quad (3.3.6)$$

Where

$$X\beta_t = (1, Z_t - 1, \dots, Z_t - q, |Z_t - 1| - E|Z_t|, \dots, |Z_t - q| - E|Z_t|, \ln h_t - 1, \dots, \ln h_t - p) \quad (3.3.7)$$

It is observed that parameter estimation implies a number of recursions, and starting-values for parameters are therefore necessary. Nelson(1991) discussed the role of starting values and concluded that in his situations the use of other starting values than the unconditional mean of $\ln h_t$ very rapidly led to values of h_t obtained by starting from the estimate of $E \ln h_t$

Chapter 4

EMPIRICAL STUDY

After exploring the general theory of ARCH-type models in the previous chapters, this section delved on fitting the GARCH and EGARCH models to the inflation series data for Kenya ranging from January 1990 to December 2015. The data was obtained from Kenya National Bureau of Statistics which is a department of the Ministry of planning and whose main role is to provide up to date information which is crucial for government planning. A description of the data is given in table 4.1.

4.1 Source of data

The study focused on CPI data sampled from all the 47 counties in Kenya. The data employed in this study comprises of 312 monthly Kenya inflation data observation from January 1990 to December 2015. Approval were obtained from the Kenya National Bureau of Statistics on the permission to use the dataset.

4.2 Analysis Methods

Preliminary analysis was first conducted to check whether the data exhibit characteristics of financial time-series data. To check for normality of the data, Jarque Bera test was conducted on the return series. Similarly, the data was examined for stationarity using ADF test. In order to establish the presence of serial correlation and the presence of ARCH effects on the data, Box-Pierce-Ljung test and Engle ARCH test was carried out respectively. This was done for both return squared return series. Model selection was based on the outcome analysis of Akaike Information Criteria and Bayesian Information Criteria(BIC). The best model was selected based on the model which had the minimum values of AIC and BIC. The rational was to obtain a parsimonious model that captures as much as possible variation in the data. After optimal model was selected, model evaluation was conducted on the selected model to establish if the selected model is adequate and meets the standard criteria for optimal model. The model should not display serial correlation, conditional heteroscedasticity or any type of non-linear dependence(Zivot,2008). Ljung Box statistics was used to test the null hypothesis of no auto-correlation up to a specific lag. Engle LM statistics was used to test null hypothesis of no remaining ARCH effects. In addition, a normal QQ plot of the standard residuals was fitted into the data. This was done in order to check whether the model has a good fit to the data and the residuals follow a white noise process.

Data Description

Descriptive and time series statistical techniques was carried out in order to provide more information regarding the data. Jarque Bera test for normality for the inflation and return

series was conducted in order to check the normality of the dataset. The test had a significant p value which means the null hypothesis of normality is rejected. The data does not follow a normal distribution which is line with characteristic of financial time series data. From table 4.1 below, the value of ADF for both the inflation-raw data and the return series has a p-value of 0.02 and 0.01 respectively. This is statistically significant and is an indication that the inflation return series is stationary. The standard deviation of the inflation data is 10.88 meaning there was a huge variation in inflation rate. Table 4.1 below gives summary statistics for the data.

Table 4.1: Data description of inflation and return series

Statistics	Inflation	Return
Mean	11.69533	-0.01638661
Standard Deviation	10.88689	0.3863861
Kurtosis	5.785278	23.66334
Skewness	2.265754	-2.626791
Maximum	61.54215	1.654138
Minimum	-3.662444	-3.324236
Jarque Bera P value	2.20E-16	2.20E-16
ADF	-3.7712	-6.1122
ADF P value	0.02082	0.01

Transformation

As stated earlier in chapter 2, variances of most financial time series changes with time and thus are non-stationary. In order to generate a stationary series, we convert the inflation

prices to returns by logarithmic transformation. From fig 4.1 above, the return series is stationary.

4.3 Empirical Study

4.3.1 Testing for Serial Correlation and ARCH effects in the return series

It is important to establish the presence of ARCH effects in the data. A time series data exhibiting conditional heteroscedasticity or autocorrelation in the squared series is said to have Autoregressive conditional heteroskedastic effects. This section uses statistical test to establish the presence of ARCH effects. Box-Pierce-Ljung test and Engel ARCH test has been employed for the purposes of testing the presence of serial correlation in the data. Box Pierce-Ljung was developed by Box and Pierce (1970) and modified by Ljung and Box (1978).

Results from table 4.2 and 4.3 below shows the presence of ARCH effects in the return inflation data. The same is also exhibited in the squared return of inflation data as shown in table and below. It is apparent that there is bunching in the variance of volatility. It means that periods of high volatility are followed by periods of high volatility while periods of low volatility and followed by periods of low volatility. This is a common characteristic of financial time series data. So the data under study exhibit the presence of ARCH effects. If the time series is an outcome of a completely random phenomenon then autocorrelation

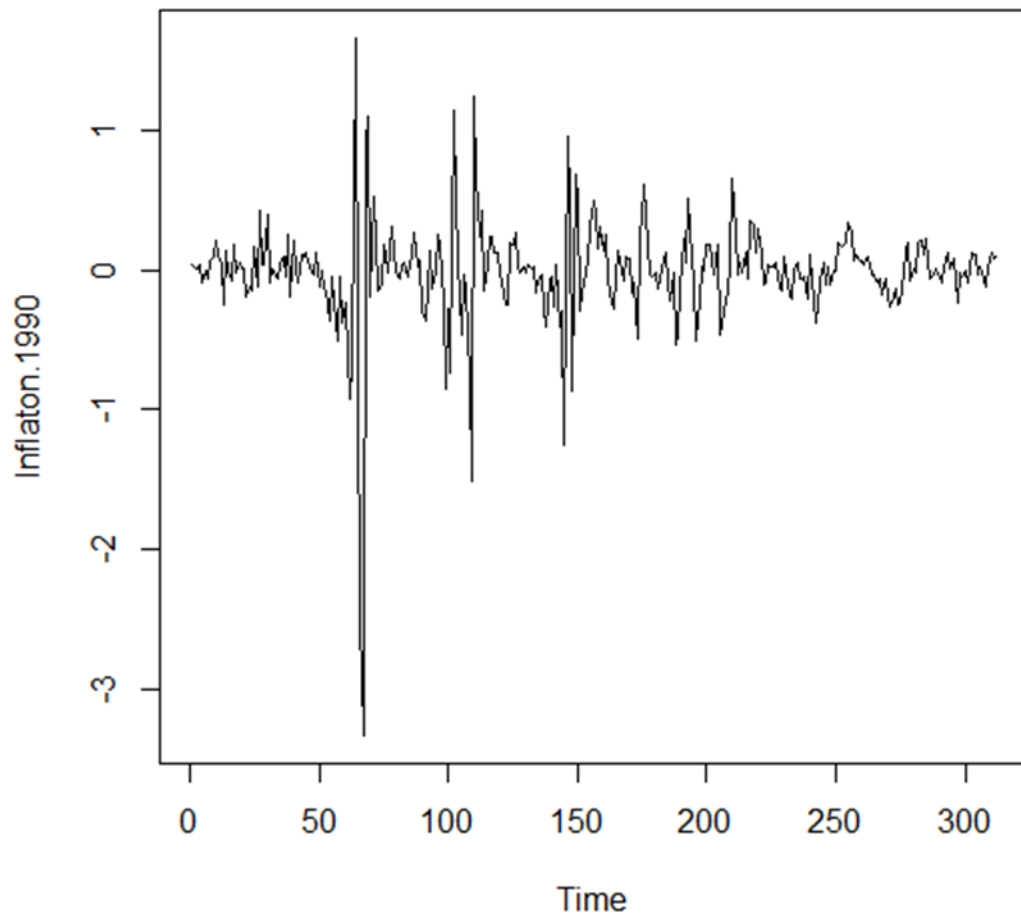


Figure 4.1: Graph of inflation return series

should be zero for all the time lags. Otherwise one or more of the autocorrelations will be significantly non-zero. The results from Ljung Box Q-test in table 4.3 and 4.4 below shows the presences of autocorrelations in the data since the p-value are significant at 95

Table 4.2: Engle ARCH test for heteroscedasticity for return series

Lag	Chisq	Df	P-value
10	88.39	10	1.21E-14
15	87.327	15	3.118E-12
20	85.935	20	3.77E-10

Table 4.3: Engle ARCH test for squared for return series

Lag	Chisq	Df	P-value
10	29.654	10	0.0009757
15	29.175	15	0.01527
20	28.699	20	0.09386

Table 4.4: Ljung Box-Q-test for returns

Lag	Chisq	Df	P-value
10	40.441	10	0.00001417
15	52.542	15	0.00000459
20	54.134	20	0.00005526

Table 4.5: Ljung Box-Q-test for squared returns

Lag	Chisq	Df	P-value
10	111.04	10	2.20E-16
15	111.84	15	2.20E-16
20	113.42	20	4.66E-15

Serial dependency in the data can also be examined through partial autocorrelation. Partial autocorrelation is a measure of the degree of association between X_t and its k th lags, X_{t-k} when the effects of other time lags $X_{t-1}, X_{t-2}, X_{t-3} \dots X_{t-k-1}$ is removed. The partial autocorrelation control for the lags. The results of partial autocorrelation in figure 4.2 below of the partial autocorrelation function graph of the return series shows that there is dependency in the squared return series. From the results it is clear that the variance of the return series is dependent on its past variance and may change with time. Overall the above result indicates that the inflation return series data is an ideal case of GARCH and EGARCH treatment.

4.3.2 Model estimation and evaluation

Model selection and analysis

The study has incorporated different methods for evaluating and selecting the best model. Appropriate models are selected based on the Akaike information Criteria (AIC) and Bayesian Information Criteria (BIC). R statistical software is used to perform trial and error tests in

autocorrelation function of return series.PNG autocorrelation function of return series.PNG

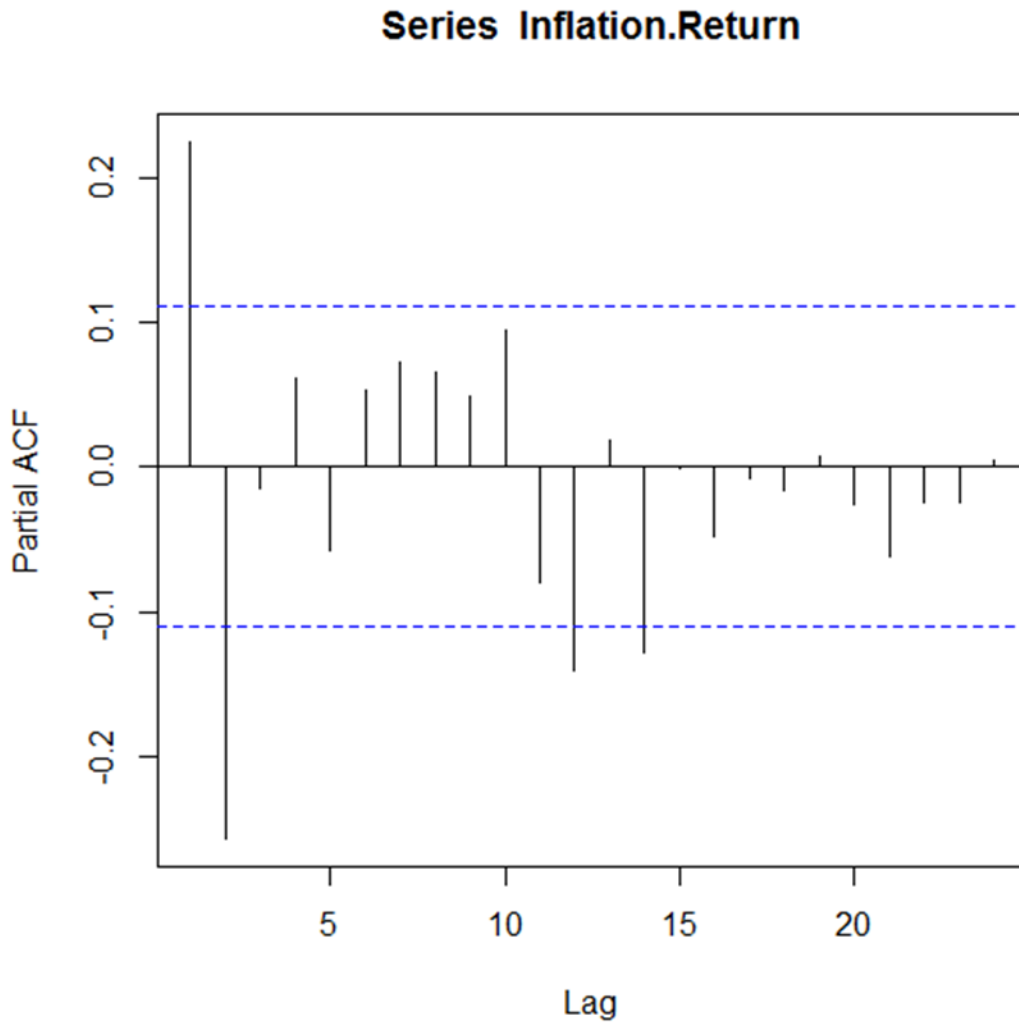


Figure 4.2: Partial autocorrelation function of return series

order to determine the best fitting model. The rational is to obtain a parsimonious model that captures as much variation in the data as possible. According to (Bollerslev, Chou and Kroner, 1992), GARCH(1; 1); GARCH(2; 1) or GARCH(1; 2) models are adequate for modelling volatilities even over long sample periods. However for the purposes of this study we have included Egarch (1, 1) in order to cater for possibility of asymmetry in the dataset. The best model should have smaller AIC and BIC values. Models with larger AIC and BIC values are considered unsuitable. Egarch model has largest negative values of the two methods adopted as the selection criteria i.e. the smallest AIC and BIC values of the fitted models as shown in table 4.6 below.

Table 4.6: Comparison of Garch and Egarch Models

Model	AIC	BIC	Log Likelihood
Garch 11	-0.11252977	-0.06465496	21.61091
Garch 12	-0.09852271	-0.0386792	20.4188
Garch 21	-0.10550257	-0.04565907	21.51115
Garch 22	1.124977	1.141961	-1104.352
Egarch 11	-0.16682	-0.10698	31.10721

Table 4.7: Parameter Estimates for EGARCH Model

Description	mu	Omega	Alpha 1	Beta 1	Gamma 1
Estimate	-0.005977	-0.465342	-0.285731	0.835255	0.694942
Standard error	0.012135	0.165281	0.07305	0.050519	0.142583
t-value	-0.49254	-2.81547	-3.91146	16.53352	4.87393
Pr(> t)	0.622335	0.004871	0.000092	0	0.000001

$$r_t = -0.005977 + \Sigma_t \tag{4.3.1}$$

$$\delta^2 = -0.465342 + 0.835255r_{t-1}^2 - 0.285731\delta_{t-1}^2 \tag{4.3.2}$$

The model fit was done based on the 95 percent confidence interval. From table 4.7 above, the p-value for alpha 1 and Beta1 are significant with small corresponding standard errors. The coefficients are thus significantly different from zero. At the same time, the standard errors which are used to measure accuracy of the model are small. This is an indication that the model produces precise estimates. It is also important to note that the estimated values of the parameters alpha1 and beta 1 satisfy the stationarity condition. Alpha1+Beta1 less than 1, with Beta1 greater than alpha1.

Table 4.8: Parameter Estimates for GARCH(1,1) Model

Description	mu	Omega	Alpha 1	Beta 1
Estimate	0.011777	0.006411	0.681349	0.406748
Standard error	0.009944	0.001914	0.12376	0.057177
t-value	1.184	3.349	5.505	7.114
Pr(> t)	0.236265	0.000812	3.68E-08	1.13E-12

From table 4.8 above, the standard errors for the estimates alpha 1 and beta 1 are small signifying precise estimates. The parameters alpha 1 and Beta 1 are also statistically significant at 95 percent confidence interval. However the stationary condition for the parameters alpha 1 and Beta 1 are not met. Alpha 1 + Beta 1=1.088097 which does not satisfy the

equation $\alpha_1 + \beta_1 < 1$.

Table 4.9: Parameter Estimates for Garch(1,2) Model

Description	mu	Omega	Alpha 1	Beta 1	Beta2
Estimate	1.19E-02	6.53E-03	6.73E-01	4.08E-01	1.00E-08
Standard error	9.26E-03	N/A	N/A	N/A	N/A
t-value	1.282	N/A	N/A	N/A	N/A
Pr(> t)	0.2	N/A	N/A	N/A	N/A

The second model to be considered is Garch (1, 2). The model satisfies the stability condition since $\alpha_1 + \beta_1 < 1$ as shown in table 4.9 above. However, the coefficients of the models are not statistically significant. Only the standard error of the mean is indicated while the standard errors for the parameters are not applicable. The result of the parameters outputs shows that the model is not appropriate for the data.

Table 4.10: Parameter Estimates for Garch(1,2) Model

Description	mu	Omega	Alpha 1	Alpha 2	Beta 1
Estimate	0.008217	0.007183	0.553598	0.25	0.310507
Standard error	0.010263	0.002149	0.130965	0.17	0.082662
t-value	0.801	3.343	4.227	1.47	3.756
Pr greater than t	0.423324	0.000828	2.37E-05	0.14	0.000172

Garch (2, 1) has relatively small AIC and BIC values. i.e AIC = -0.10550257 and BIC = -0.04565907 as indicated in table 4.10. It also has small standard errors and α_1 and

Beta 1 are statistically significant at 95 percent confidence interval except Alpha 2 which is not statistically significant hence the model is not a good fit to the data. In addition, the parameters violate the stationarity condition. $\text{Alpha 1} + \text{Alpha 2} + \text{Beta 1} = 1.115037$.

Table 4.11: Parameter Estimates for Garch(2,2) Model

Description	mu	Omega	Alpha 1	Alpha 2	Beta 1	Beta 2
Estimate	-5.04E-03	1.13E-02	1.68E-01	1.00E-08	4.90E-01	2.97E-01
Standard error	8.51E-03	NA	2.75E-02	NA	NA	NA
t-value	-0.593	NA	6.121	NA	NA	NA
Pr greater than t	0.554	NA	9.30E-10	NA	NA	NA

The final model to be considered is Garch (2, 2). The model does not fit the data well as it has the largest AIC and BIC values as compared to other models fitted. Additionally from table 2.2, the p-values of coefficient parameters Alpha1, Beta 1 and Beta 2 are not statistically significant at 95 percent confidence interval. The fact that the p-values for the coefficient parameters as well as standard errors are not applicable is an indication that the model is a poor fit to the data, hence cannot be used for forecasting future inflation. In conclusion, it has been established that Egarch(1,1) is the best fit model among the Garch and Egarch models for the Kenya inflation data as it satisfies all the criteria for selection

4.3.2.1 Evaluation of estimated EGARCH (1, 1) Model

After the EGARCH model had been fit to the data, the adequacy of the fit was evaluated using various graphical and statistical diagnostics techniques. If the Garch model is correctly

specified, then the estimated standardized residuals should not display serial correlation, conditional heteroscedasticity or any type of non-linear dependence (Zivot, 2008). Ljung Box statistics has been used to test the null hypothesis of no autocorrelation up to a specified lag. And Engles LM statistic was also used to test null hypothesis of no remaining ARCH effects. This is clearly the case as shown in table 4.12 below which shows that there is no presence of correlation in the data. In addition, table 4.13 shows the absence of ARCH effects in the data. One of the assumptions of ARCH-type model is that, for a good fit model, the residuals must follow a white noise process; the residuals are expected to be random; independent and identically distributed following the normal distribution. A normal QQ-plot of the standard innovations or residuals of the fitted model should look roughly linear. The QQ-plot was conducted for the fitted model as shown in figure 4.3 below and the graph look linear as majority of the points are on the line and hence proves the model to be normal fit for the data. Further a plot of residual against time as shown in fig 4.5 is stable about zero hence stationary and shows the presence of volatility clustering which is characteristic of financial time series data. To check for the absence of autocorrelation, further analysis of the graph of ACF of standard residuals was conducted. The graph in fig 4.4 clearly shows the absence in autocorrelation of the squared residual. The diagnostic tests conducted proves that EGARCH(1,1) model is suitable for inflation series data and therefore can be used for forecasting future inflation

Table 4.12: Weighted Ljung-Box Test on Standardized Squared Residuals

Lag	Statistic	P-value
Lag[1]	0.003702	0.9515
Lag[2*(p+q)+(p+q)-1][5]	2.168663	0.5791
Lag[4*(p+q)+(p+q)-1][9]	6.067096	0.2906

Ho: No serial correlation

H1: No serial correlation

From the table 4.12 above there is no serial correlation exhibited by the model as the p-values are not significant at 95 percent confidence interval

Table 4.13: Weighted ARCH LM Tests

Lag	Statistic	Shape	Scale	P-value
ARCH Lag[3]	2.296	0.5	2	0.12974
ARCH Lag [5]	3.636	1.44	1.67	0.20992
ARCH Lag[7]	6.879	2.32	1.54	0.09221

Ho: No presence of ARCH-effects

H1: There is presence of ARCH effects

Table 4.13 above shows that there is not presence of ARCH effect as the p-values are not significant at 95 percent confidence interval

Plot.PNG Plot.PNG

norm - QQ Plot

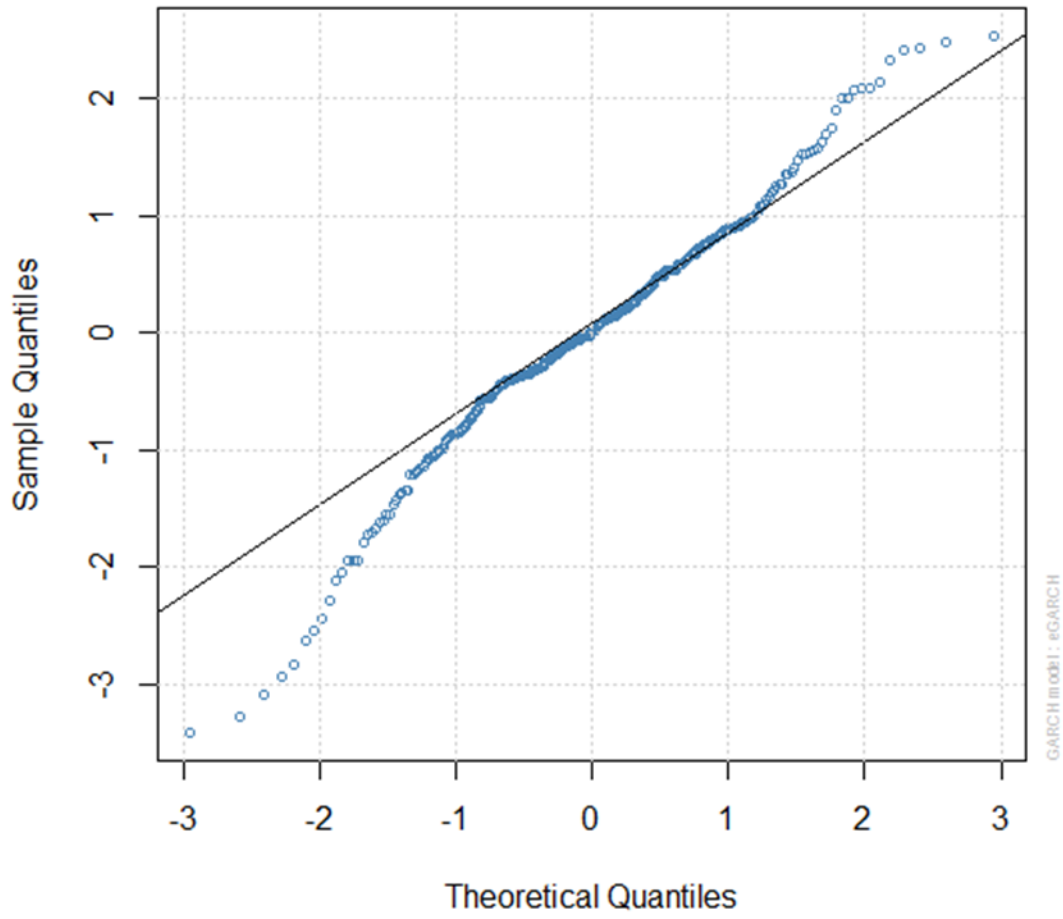


Figure 4.3: QQ Plot for normality

of Squared Standard Residuals.PNG of Squared Standard Residuals.PNG

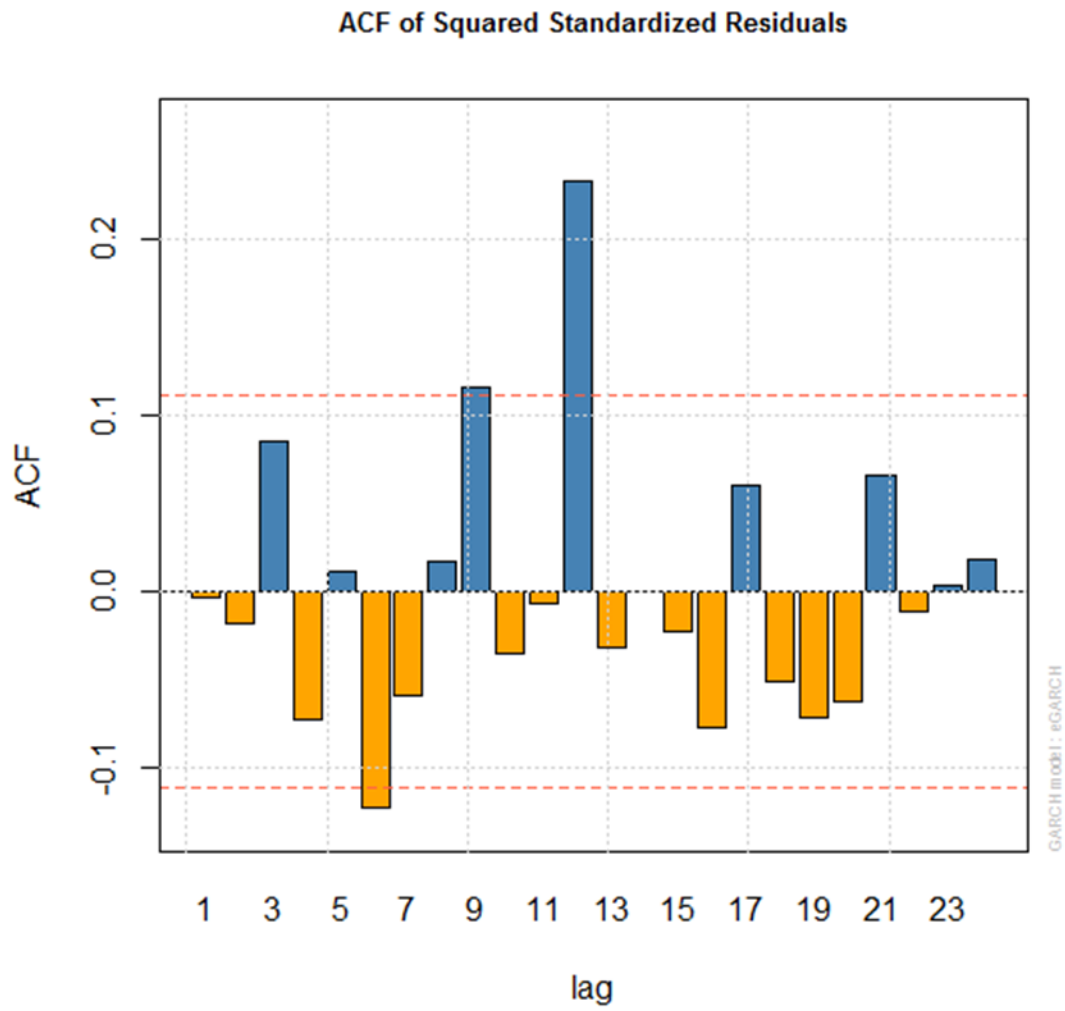


Figure 4.4: ACF of Squared Standard Residuals

of residual.PNG of residual.PNG

res

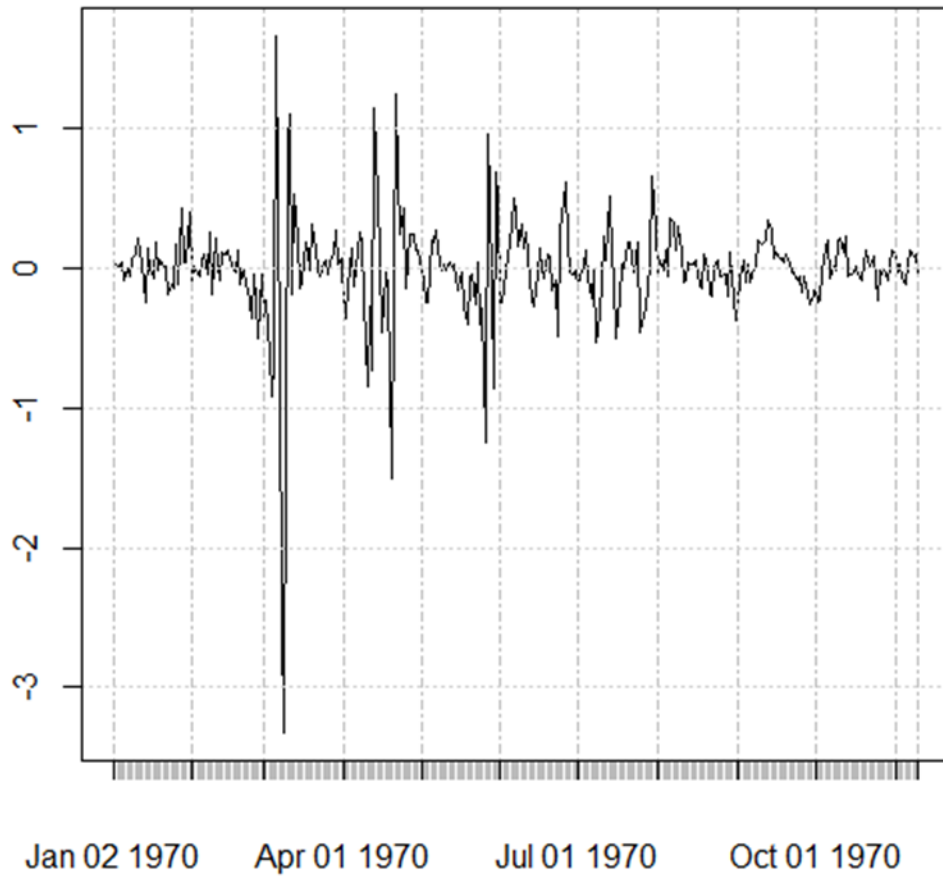


Figure 4.5: Plot of residual against time

Forecasting using EGARCH Model

After model evaluation and diagnostics, out of sample forecast for EGARCH model was conducted using the inflation data. The table below summarizes the forecast values for the next one year.

Table 4.14: Forecast Inflation Data

Time	Forecast Inflation
T+1	0.1359
T+2	0.1499
T+3	0.1624
T+4	0.1733
T+5	0.1826
T+6	0.1905
T+7	0.1972
T+8	0.2027
T+9	0.2074
T+10	0.2112
T+11	0.2143
T+12	0.2169

Chapter 5

DISCUSSIONS, CONCLUSION AND RECOMMENDATION

5.1 Discussion

From the graph of the return series in Fig 4.1 , the inflation return series shows series of spikes which is an indication of volatility and subsequent volatility clustering. This is a very key feature of financial time series and is not surprising in the case of Kenya inflation. EGARCH(1,1) has been fitted and the parameters α_1 and β_1 which are -0.285731 and 0.835255 are all statistically significant. The standard errors which are used to measure the accuracy of the model is small indicating better accuracy of the model fit to the data. It is also an indication that the model produces precise estimates. In addition, the estimated values of the parameters of the model α_1 and β_1 satisfy the stationary conditions. In this regard, it is important to point out that GARCH(1,1) was the closest model in terms of minimum AIC and BIC to the EGARCH(1,1) model. However, GARCH(1,1) model does not meet the stationary condition which as the parameters α_1 and β_1 are greater than 1. Diagnostic tests

for EGARCH(1,1) model was conducted in terms of ARCH-LM tests, Ljung BOX test and QQ normality plot. The results shows the absence of ARCH effect. The Ljung box test also shows no presence of serial correlation.

5.2 Conclusion

The importance of a model for forecasting future inflation for a given country cannot be overstated. From the analysis generated from this study, it has emerged that EGARCH model is the best fit model for the Kenyan inflation data. EGARCH(1,1) model has the smallest AIC and BIC compared to various GARCH models. The AIC and BIC criteria are used to determine the best model. It is also noteworthy to state that the model has satisfied all the conditions which include stationary conditions, absence of ARCH effect and random and independently distributed residuals. Further more, the model exhibit a normal distribution as exhibited in figure 4.5 of the QQ plot. The study has presented us with the opportunity to have a deeper understanding of the theory of time series analysis and its application to real life situation. It is clear from the study that the Kenyan inflation rate data series is characterized with spikes, variations and trends, hence EGARCH model serves as the best in forecasting Kenyan inflation data. Other areas of application include environmental and pollution data (Peng and Dominici,2008).

5.3 Recommendation

The study would recommend further research on modeling Kenyan inflation data using a multivariate Arch-type models where other variables that have influence on inflation are incorporated. The study would also recommend further research on modeling of performance of Nairobi Stock Exchange using EGARCH model.

References

Ahiati, V.S (2007), Discrete Time Series Analysis with ARMA Models, Revised Edition, Holden-Day, Oakland.

Amos, C (2010), Time Series Modelling with Application to South African Inflation Data, A thesis submitted to University of Kwazulu Natal for the degree of Masters of Science in Statistics.

Awartani, B.M.A and Corradi, V. (2005) predicating the volatility of the S and P-500 stock index via GARCH models: The role of asymmetries. *International Journal of Forecasting*, 21, 167-183

Baillie, R.T and T. Bollerslev (1992). Prediction in Dynamic Models with Time Dependent Conditional Variances. *Journal of Econometrics*, Vol.52, No.1, pp.91-113

Bailey, J. M. (1956). The Welfare Cost of Inflationary Finance. *Journal of political economy*, 64, 93-110.

Bera, A. and Higgins, M.I. (1993) Arch models: properties, estimation and testing. *Journal of Economics Surveys*, 7, 307-366

Bollerslev, T (1986), Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31, 307-327.

Bowerman, B. L., and O'Connell, R. T. (1997). Applied Statistics: Improving Business Processes. Irwin Series, Times Mirror Higher Education Group, Inc. Chicago.

Brooks, C (2008), Introductory Econometrics for Finance, 2nd ed. Cambridge University Press, New York.

Chateld, C (2000), Time Series Forecasting: University of Bath, Chapman and Hall text in statistical science, London

Campbell, J.Y, Lo, A.W. and Mackinlay, A.C (1997), Modelling daily value at risk using realized volatility and ARCH type models, Journal of empirical nance volume 11, no. 3 pages 379-398.

Diggle, P, Hearnerty, P and Liang, K.Y (2002) Analysis of longitudinal data, Oxford University press.

Engle, R.F., 1983, Estimates of the variance of U.S. inflation based on the ARCH model, Journal of Money Credit and Banking 15, 286-301.

Engle, R (1982), The Use of ARCH/GARCH Models in Applied Econometrics; Journal of Economic Perspectives-Volume 15, Number 4. pp. 157 - 168.

Engle, R. F. and Bollerslev, T. Modeling the Persistence of Conditional Variances, Econometric Review, 5, 150.1986

Garcia, R.C., Javier, C., Marko, A., Garcia, B.S., (2003), A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices; A paper submitted to the IEE power system Journal in Berlin.

Goudarzi, H., and Ramanarayanan, C. S. (2010). Modeling and estimation of volatility in

the Indian stock market. *International Journal of Business and Management*, 5(2), 85-98.

Hall, R., (1982), *Inflation, Causes and Effects*, Chicago University Press, Chicago

Jean - Philippe, P. (2001). Estimating and Forecasting Volatility of Stock indices using Asymmetric GARCH models and (Skewed) Student-t densities. *Ecole d'Administration des Affaires*, University of Liege, Belgium.

Lastrapes, W.D. (1989) Exchange rate volatility and US monetary policy: An ARCH application. *Journal of money, credit and banking*, 21, 66-77

Lopez, J.A (1999), Evaluating the predictive accuracy of volatility models, *Economic research*, Department of federal reserve bank of San Francisco.

Mandelbrot, B. (1963) The variation of certain speculative prices. *Journal of Business*, 36, 394-419

Minkah, R. (2007). *Forecasting Volatility*. (Unpublished Masters thesis). Department of Mathematics, Uppsala University, Uppsala, Sweden.

Moser, G., Rumler, F. and Scharler, J. (2004.) "Forecasting Austrian Inflation," Working Papers 91, Oesterreichische Nationalbank (Austrian Central Bank).

Ngailo, E. (2011). *Modelling and Forecasting using time series GARCH models: An Application of Tanzania inflation rate data*. (Unpublished Masters Thesis). University of Dar es Salaam.

Nelson, D. B. (1991). *Conditional Heteroscedasticity in Asset returns: A New Approach*,

Econometrica, 59 , 347 -370.

Nelson, D. B. and Cao, C. Q., 1992, Inequality constraints in the univariate GARCH model. Journal of Business and Economic Statistics, 10: 229-235.

Peng R.D and Dominici F (2008) Statistical methods for Environmental Epidemiology with R Springer Series in Statistics, New York.

Peng, R. D., and Dominici, F. (2008). Statistical methods for environmental epidemiology with R. *R: a case study in air pollution and health*.

Perniaag, F.E., Nor, M., Ekonomi, P.P (2004), Modelling and Forecasting on the Malaysian Inflation Rates: An Application of GARCH Models, Bangi, UKMpress.

Su, C. (2010). Application of EGARCH Model to Estimate Financial Volatility of Daily Returns: The empirical case of China. (Unpublished masters thesis) University of Gothenburg, Sweden.

Shephard, N. (1996). Statistical Aspects of ARCH and Stochastic Volatility Time Series in Econometrics Finance and other fields, Chapman and Hall, London.

Shamiri A. and Isa, Z. (2009) Modelling and forecasting volatilities of the Malaysian stock market. Journal of mathematics and statistics, 5,234-240

Sobel RS, Stroup RL, Macpherson DA, Gwartney JD (2006) Understanding Economics. Thompson South-Western, Mason, USA, p 343

Tsay, R.S., 2002. Analysis of Financial Time Series". ISBN 0-471-41544-8. Webster, D.,

2000. Webster's New Universal Unabridged Dictionary. Barnes and Noble Books, New York.

Tsay, R.S. (2010) Analysis of financial time series. New Jersey: John Wiley and Sons

Talke, I. S (2003), Modelling volatility in time series data, MSc thesis, University of Kwa-Zulu Natal.

Tsay, R. S. (2002). Analysis of Financial Time Series. 2nd Edition. New York, John Wiley and Sons, Inc

Weiss, A. A. (1984). ARMA models with ARCH Errors. Journal of Time Series Analysis, 5, 129 - 143.

Webster, D (2000), Websters New Universal Unabridged Dictionary. Barnes and Noble Books, New York.

Yuksel, H.,and Bayram, H. (2005). ARCH - GARCH modelling in Turkish, Greek and Russian Stock Markets. Seminar in Financial Data Analysis

Zakaria,S. and Abdalla, S. (2012) modelling stock return volatility: Empirical evidence from Saudi stock exchange. International Research Journal of Finance and Economics, 85,166-179