Modelling Inflation Rate Volatility in Kenya Using Arch -Type Model Family

Johnson Okeyo[1] Mwaniki Ivivi[2] Philip Ngare[2]
1.School of Mathematics, University of Nairobi, Kenya

Abstract
This paper describe the empirical study based on financial time series modelling with special application to modelling inflation data for Kenya. Specifically the theory of time series is modelled and applied to the inflation data spanning from January 1985 to April 2016 obtained from the Kenya National Bureau of Statistics. Three Autoregressive Conditional Heteroscedastic (ARCH) family type models (traditional ARCH, Generalized ARCH (GARCH), GJR GARCH and the Exponential GARCH (EGARCH)) models were fitted and forecast to the data. This was principally because the data were characterized by changing mean and variance. The outcome of the study revealed that the ARCH –family type models, particularly, the EGARCH (1, 1) with generalized error distribution (GED) was the best in modelling and forecasting Kenya's monthly rates of inflation. The study recommends that governments, policy makers interested in modelling and forecasting monthly rates of inflation should take into consideration Heteroscedastic models since it captures the volatilities in the monthly rates of inflation.
Keywords: Inflation, Volatility, GARCH

1. Introduction

The performance of every financial sector in an economy is affected by a foreseeable rise in the inflation rate which interferes with the process of allocating and distributing resources effectively. The stability of the price and unwavering inflation is a major objectives of every government because it is an significant financial indicator that the economists, governments, political leaders, shareholders, stakeholders and investors use as their foundation of argument when discussing on the state of the economy.

According to Webster (2000), a continuous increase of the prices of commodities in an economy is referred to us as inflation. Whenever an economy experiences an increase in the rate of inflation the purchasing power of the local currency decreases and therefore an individual is not able to buy as much commodities compared to when the inflation rate is stable. Financial analysts conclude therefore that there exist direct and indirect aftershocks of inflation in every sector of the economy ranging from inflation rates, investment, unemployment, interest rates, and stock markets among others.

An increase in the inflation rate has been found in most of the recent studies to affect financial institutions and investments leading to negative effects .Rising inflation has resulted into a major economic challenges facing most countries in the world especially developing countries like Kenya. For the success of every economic policy inflation rate should be considered as a major focus for growth and development.

A research study performed by Ngailo and Massawe (2014) on the modelling inflation rate in Tanzania using ARCH family models. The data taken was from 1997 to 2010. The goodness of fit was assessed through the Akaike information criteria (AIC), Bayesian information criteria (BIC) and standard error (SE) Based on minimum AIC and BIC values, the best fit GARCH models tend to be GARCH (1; 1) and GARCH (1; 2) After estimation of the parameters of selected models, a series of diagnostic and forecast accuracy test were performed. GARCH (1,1) model was found to be the best fitting model and was used for forecasting . it was concluded that the actual series were close to the forecasted series.

Numerous evaluation criteria were established to measure the performance of these models. One criterion that is popularly used is by estimating the maximum likelihood of the models and observing which of them has the highest log-likelihood value (Shephard, 1996). In situations when the models do not have the same number of parameters, the principle of parsimony is applied and a suitable model selection criterion such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and the Hannan–Quinn (HQ) is used to choose the best model. Another model criterion is to submit estimated models to misspecification tests and observe each model perform under each scenario.

The estimated models were checked to verify if it adequately represents the series. Diagnostic checks were performed on the residuals to see the validity of the distribution assumption. The (Q-Q) plot, skewness and kurtosis of the residuals were used to check for the validity of the distribution assumption. R programming software version 3.3.1 was used for data analysis.

The objective of this study

The general objective of the study was modelling monthly inflation rates in Kenya using the ARCH, GARCH, GJR GARCH and EGARCH models and determines the best among these models in forecasting the inflation rate returns in Kenya. The specific objectives of the study were:

i. To fit the ARCH type models to the data.
ii. To analyze the adequacy of the fitted E-GARCH(1,1) model.
iii. To forecast the fitted E-GARCH (1,1) model.

The idea of this study is to come up with a better parsimonious model among the ARCH family models which captures the dynamic nature of the inflation data. The best model found for the study is E-GARCH(1,1) with Generalized error distribution which gives the best asymmetry and leverage effect. GED was found to be the best fitting since it makes better fitting estimation of the parameters relatively easier.

The paper proceeds as follows: section two presents the proposed empirical models, section three discusses empirical findings and their analysis, together with a brief summary and conclusions drawn in section four.

2. METHODOLOGY

Various GARCH, GJR-GARCH & EGARCH models were considered for modelling the inflation rate return series and a model was selected based on the one that had the least AIC, the highest Log likelihood and the QQ plot.

2.1: The ARCH Model

The ARCH models were introduced by Engle (1982). They are used to model financial time series data. Suppose $Y_1, Y_2, \ldots, Y_t$ are the time series observations and let ψ_t be the set of y_t up to time t, including y_t for $t \le 0$. The process $\{y_t\}$ is an Autoregressive Conditional Heteroscedastic process of order p, ARCH(p), if:

$$r_t = \mu + y_t$$

$$y_t = \sigma_t \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma^2) \tag{3.1}$$

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j y_{t-j}^2 \tag{3.2}$$

With $\alpha_0 \ge 0,\ \alpha_j \ge 0$ and $\sum_{j=1}^{p} \alpha_j < 1$, as the ARCH model parameter limits. The conditions stated guarantees that the conditional variance be positive

2.2: The GARCH Model

The process Y_t is a Generalized Autoregressive Conditional Heteroscedastic model of order p and q, GARCH (p,q) if:

$$r_t = \mu + y_t$$

$$y_t = \sigma_t \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma^2)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \ldots\ldots + \alpha_p y_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \ldots\ldots. + \beta_q \sigma_{t-q}^2$$

$$= \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{3.3}$$

Where q>0, $p \ge o, \alpha_0 > 0, \alpha_i \ge 0$ for i=1,2,…….,p, $\beta_j \ge 0$ for j=1,……..q are the GARCH model parameter limits. Again these conditions are needed to guarantee that the conditional variance $\sigma_t^2 > 0$.

The Kurtosis is given by;

$$K = \frac{E(y_t^4)}{\{E(y_t^2)\}^2} \tag{3.4}$$

Substituting equation (3.15) and equation (3.17), we get;

$$K = 3\frac{1-(\alpha_1+\beta_1)^2}{1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2} \tag{3.5}$$

Which is strictly greater than 3 unless $\alpha_1 = 0$
The same fitting procedure is applicable for a general GARCH (p,q).

2.3 GJR-GARCH
The GJR model gives the possible asymmetries observed from the data (Brooks, 2008). Glosten, Jaganathan and Runkle (1993) developed a type of GARCH extension model that permits the conditional variance to have different reaction to the results of past negative and positive innovations.

$$\delta_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i u_{t-i}^2 + \gamma_i u_{t-i}^2 d_{t-1} + \sum_{i=1}^{p}\beta_j\delta_{t-j} \tag{3.6}$$

where d_{t-1} is a dummy variable that is:

$$d_{t-1} = \begin{cases} 1 \text{ if } u_{t-1} < 0, & \text{bad news} \\ 0 \text{ if } u_{t-1} \geq 0, & \text{good news} \end{cases}$$

In the model, effect of good news is shown by α_i, while bad news shows their impact by $\alpha + \gamma$. In addition if $\gamma \neq 0$ news impact is asymmetric and $\gamma > 0$ leverage effect exists.

For the satisfaction of non-negativity condition coefficients would be $\alpha_0 > 0$, $\alpha_i > 0$, $\beta \geq 0$ and $\alpha_i + \gamma_i \geq 0$. That is the model is still acceptable, even if $\gamma_i < 0$, provided that $\alpha_i + \gamma_i \geq 0$.

2.4 E-GARCH
Exponential GARCH (EGARCH) proposed by Nelson (1991) gives a leverage effects and asymmetry in its equation. In the EGARCH model the specification for the conditional covariance is given by the following form:

$$\ln\left(\delta^2{}_t\right) = \alpha_0 + \sum_{j=1}^{q}\beta_j \ln\left(\delta_{t-j}\right) + \sum_{i=1}^{p}\alpha_i\left|\frac{u_{t=i}}{\sqrt{\delta_{t=i}}}\right| + \sum_{k=1}^{r}\gamma_k\frac{u_{t-k}}{\sqrt{\delta_{t-k}}} \tag{3.7}$$

Two advantages stated in Brooks (2008) for the pure GARCH specification; by using $\ln\left(\delta^2{}_t\right)$ even if the parameters are negative, will be positive and asymmetries are allowed for under the EGARCH formulation.

In the equation γ_k represents leverage effects which accounts for the asymmetry of the model. While the basic GARCH model requires the restrictions the EGARCH model allows unrestricted estimation of the variance.

If $\gamma_k < 0$ it indicates leverage effect exist and if $\gamma_k \neq 0$ impact is asymmetric. The meaning of leverage effect bad news increase volatility.

2.5 Model selection criteria
Various selection criteria models like Akaike information criterion, log likelihood (LL), Bayesian information statistics (BIC), Shwartz information Criterion (SIC) and Hannan-Quin (HQ) were used but Akaike information criterion (AIC) was chosen to be the best criterion for this paper to give us the lowest minimized values and selecting a parsimonious model.

3. RESULTS ANALYSIS
3.1 Data description
Secondary data consisting of year-on-year inflation data for each month from January 1985 to April 2016 was used in this study. The total number of data points is therefore 376. The year-on-year inflation is the percentage change in the consumer price index (CPI) over a twelve-month period which is used to measure changes over time in the general price level of goods and services that households acquire for the purpose of consumption. The monthly year-on-year inflation is collected by the Kenya National Bureau of Statistics.

www.knbs.or.ke

The time plot of inflation data as shown in fig 3.1

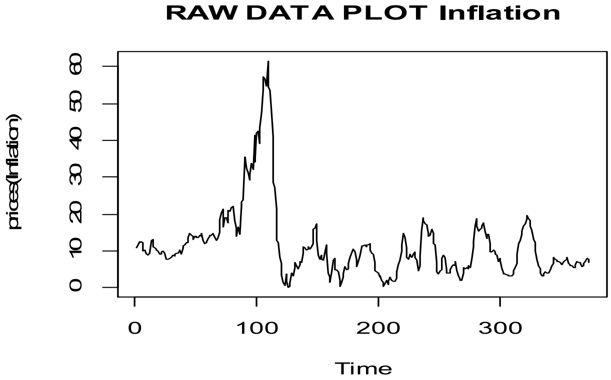

Figure 3.1: The plot of raw data

The inflation rate trend can be clearly seen observed as the rate rises gradually. However sudden and sharp rises (volatility clustering) can be seen. The volatility characteristics of financial time series data can be clearly seen from the rise drop of the inflation rate.

Table 3.1 Simple and log Returns Descriptive Statistics

	Simple Returns	Log Returns
Nobs	376	376
NAs	0	0
Minimum	-0.91801	-2.501104
Maximum	4.207659	1.65013
1 Quartile	-0.08624	-0.090191
3 Quartile	0.096058	0.09172
Mean	0.049766	-0.002058
Median	0.000506	0.000505
Sum	18.71214	-0.773666
SE Mean	0.02099	0.016332
LCL Mean	0.008494	-0.034171
UCL Mean	0.091039	0.030056
Variance	0.165655	0.100289
Stdev	0.407007	0.316685
Skewness	4.911781	-0.881051
Kurtosis	37.91698	15.20271

The resulting values indicate that the returns have a small variance and consequently a small standard deviation. Finally values of the kurtosis 37.91698 of simple returns and 15.20271 log returns are greater than 3 hence the data exhibits excess kurtosis. The resulting kurtosis and skewness suggests that the data is non stationary.

3.2 Time series returns plot

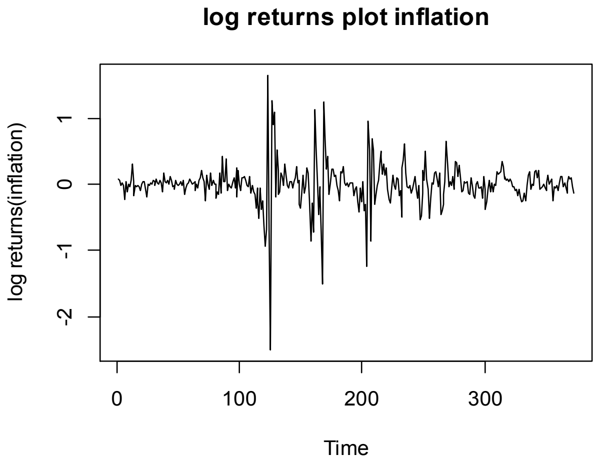

Figure 3.2 Log returns

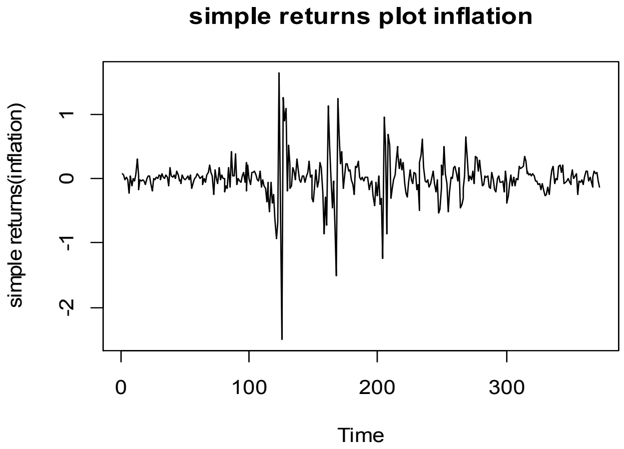

Figure 3.3 Simple returns

The time plot for Arithmetic (Simple) and Log returns is as shown in figure 3.2 and Figure 3.3The plot of the simple and log returns of inflation rate data. From the simple and log returns plots of the returns, volatility clustering can be clearly observed. (i.e where a large changes are followed by other large changes of either sign and small changes are followed by small changes).

The mean reverting (returns tend to remain around a certain value) property can also be seen clearly where the returns revolve around zero. Extreme return can also be observed.

3.3 EXPONENTIAL GARCH MODEL

E-GARCH (1,1) Normalized, Student - t and GED Error distribution QQ Plots.

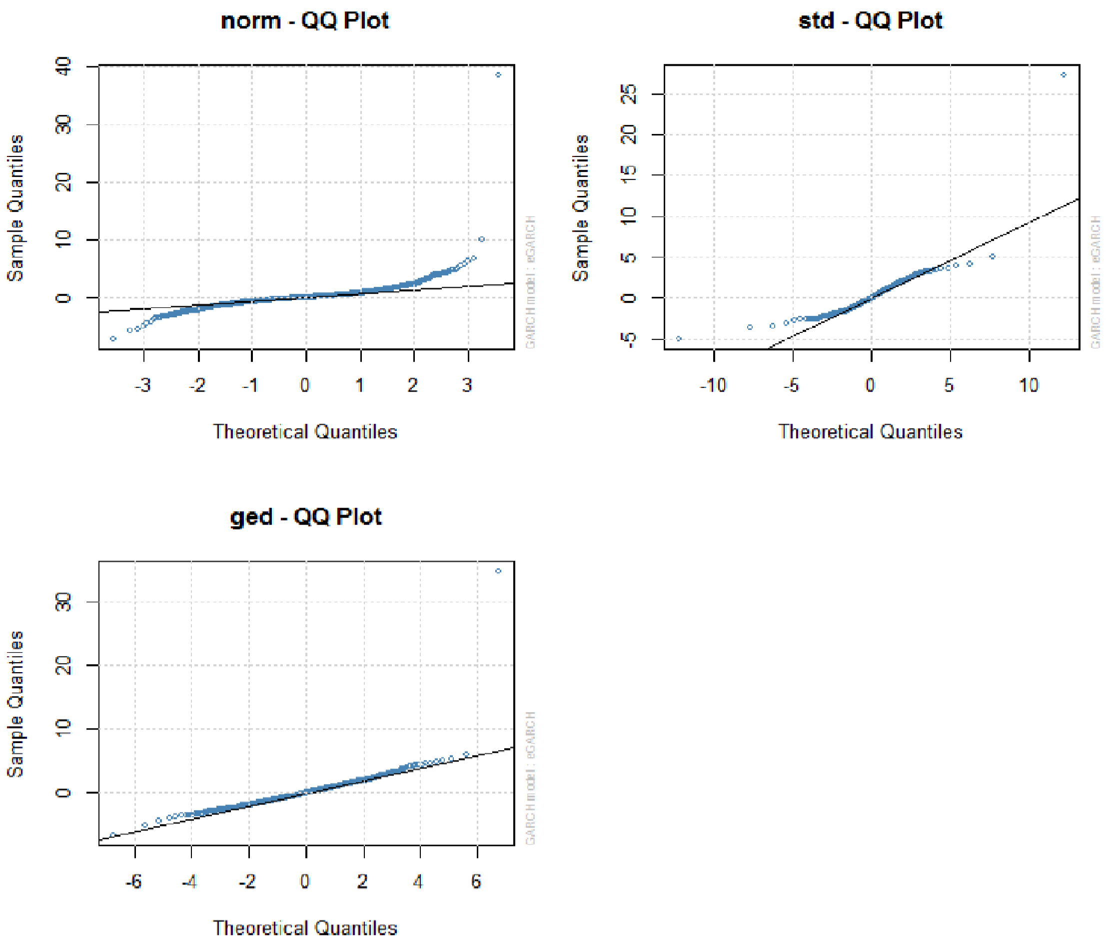

Figure 3.4 The QQ plot for E-GARCH (1,1) with normal and student-t and GED residual distribution.

The norm - QQ plot suggests that the data is a poor fit when the residual distribution is Normal, thus the data is possibly non-normal. This is clearly indicated by the failure of the data to be linear at the tails suggesting a more heavily tailed distribution for the residuals. The std – QQ plot seems to have a relatively fair fit with student-t distribution being the residual distribution. However, the ged – QQ seems to be a good fit for the data since the plot is almost linear.

3.4 MODEL COMPARISON AND SELECTION CRITERION.

Table 3.2: The comparison of the AIC, BIC, SIC, HQ and the Log Likelihood of GARCH (1,1), GJRGARCH (1,1) & E-GARCH (1,1).

MODEL	DISTRIBUTION	AIC	BIC	SIC	HQ	LL
GARCH (1,1)	GED	0.5118	0.4697	0.5121	-0.4951286	107.1767
GJR-GARCH (1,1)	GED	--0.5339	-0.4814	-0.5343	-0.51311	104.5884
E-GARCH (1,1)	**GED**	**-0.5375**	**-0.4849**	**-0.5378**	**-0.51664**	**105.2472**

Based on the results in table 3.2, the model that appears to best fit the dataset is E-GARCH (1,1) with Generalized error distribution. The asymmetric E-GARCH model is a better fit than the GARCH model since it can capture the leverage effects unlike the GARCH model. The E-GARCH model also has an advantage over the GJR-GARCH model in that, even if the parameters are negative, variance remains positive because the variable modelled is In (δ_t^2).

However, the QQ-plots, AIC suggested that E-GARCH model with generalized error distribution is a better fit for the dataset. The model selected therefore is E-GARCH (1,1) with generalized error distribution.

3.5 MODEL AND RESIDUAL ANALYSIS.

Table 3.3: E-GARCH (1,1) with GED conditional distribution parameter estimates.

Conditional Variance Dynamics				
GARCH Model	E-GARCH (1,1)			
Mean Model	ARFIMA(0,0,0)			
Distribution Model	GED			
Optimal Parameters				
Parameter	**Parameter Estimate**	**Standard Error**	**t-value**	**P-Value**
Omega	-0.32931	0.130518	-2.5231	0.011633
Alpha1	-0.22027	0.058243	-3.7819	0.000156
Beta1	0.9032	0.03875	23.3085	0.000000
Gamma1	0.5521	0.164732	3.3515	0.000804
Shape	1.20473	0.121957	9.8783	0.000000

Gamma1 is positive and statistically significant at 1% level of significance. The magnitude is 0.5521 and the sign is positive. The positive value shows that past positive events have more influence on future volatility. The positive and statistically significant coefficient shows that the inflation rate volatility is higher after positive shock i.e. bad news generates less volatility than good news.

Alpha1 and beta1 are statistically significant at 1% significance level. The sum of

$\alpha + \beta =$ 0.68293 $<$ 1. This shows that volatility is persistent.

Table 3.4: E-GARCH (1,1) with GED conditional distribution Weighted Ljung-Box Test on Standardized Residuals

Lag	**Test statistic**	**p-value**
Lag[1]	20.76	5.21E-06
Lag[2*(p+q)+(p+q)-1][2]	22.6	9.27E-07
Lag[4*(p+q)+(p+q)-1][5]	24.07	1.45E-06
H_o: No serial correlation		

Table 3.5: E-GARCH (1,1) with GED conditional distribution Weighted Ljung-Box Test on Standardized Squared Residuals

Lags	statistic	p-value
Lag[1]	0.00428	0.9478
Lag[2*(p+q)+(p+q)-1][5]	1.88127	0.6467
Lag[4*(p+q)+(p+q)-1][9]	5.9195	0.3076

The high p-values of 0.9478, 0.6467 and 0.3076 lead us to accept the null hypothesis which is further strengthened by the observation of the ACF;

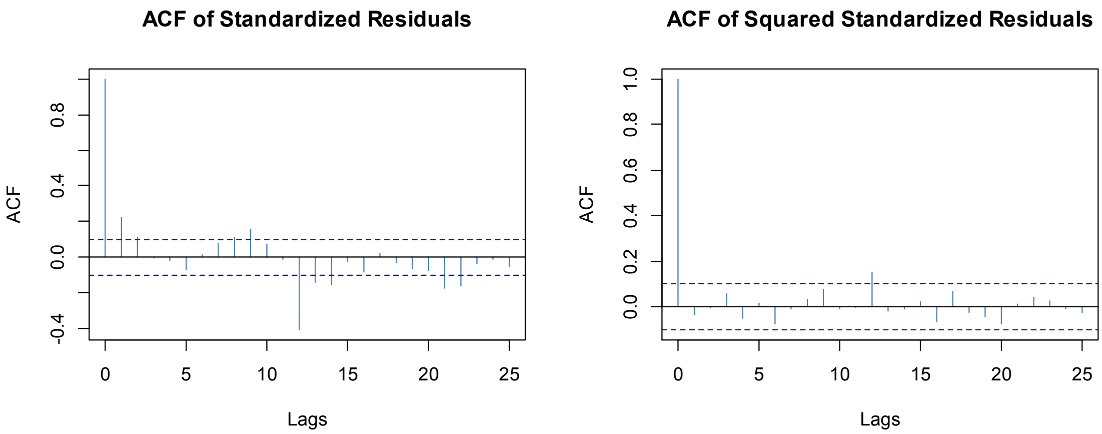

Figure 3.4: The ACF of Standardized Residuals and Squared Standardized Residuals

Table 3.6: E-GARCH (1,1) with GED conditional distribution Weighted ARCH LM Tests results.

ARCH Lag	Test Statistic	Shape	Scale	P-Value
ARCH Lag[3]	1.922	0.5	2	0.16569
ARCH Lag[5]	3.216	1.44	1.667	0.26001
ARCH Lag[7]	6.774	2.315	1.543	0.09711

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

3.6 Forecasting

In the figure 3.5 below, a volatility forecast is done for the inflation rates. As we can see, the volatility forecast converges to the unconditional variance.

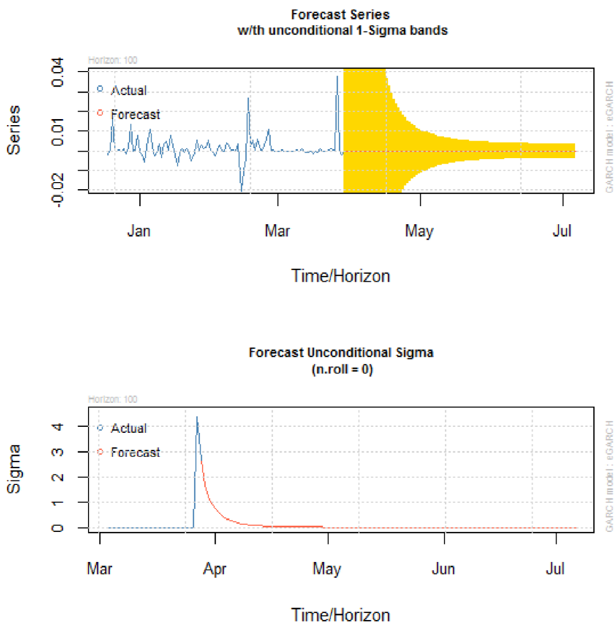

Figure 3.6: Forecast Series and Forecast Volatility.

4. Conclusion

This study opted for all specifications of the GARCH models although empirical evidence shows that the lower specifications are able to sufficiently capture the characteristics of inflation rates while at the same time upholding the principle of parsimony. Various GARCH, GJR-GARCH and E-GARCH models were fitted with variations being made to the conditional distribution used i.e. normal, student-t and generalized error distribution.

The best fitting model was selected based on the lowest AIC, BIC, SIC and the LL. The resultant model was found to be an E-GARCH (1,1) with a GED conditional distribution. The E-GARCH model provides a better fit than the GARCH model and its advantages over the GARCH model are that first, it can capture leverage effects and secondly, that there is no restriction that the parameters α_1 and β_1 must be positive. For the inflation rate data this value is positive meaning that a positive shock has more impact on inflation rate volatility than a negative shock. This is in contrast to the leverage effects results in the developed countries. The E-GARCH model was able to capture this.

Stylized facts about financial time series such as volatility clustering and persistence were also observed. Normality assumption was rejected in the original inflation rates data as well as in the residuals of the fitted model. This was inferred from the QQ-plots, skewness and kurtosis coefficients. The chosen E-GARCH(1,1) model with the generalized error distribution was better suited to accommodate the skewness and kurtosis in the inflation rate return series.

Acknowledgements
I would like to thank my supervisors Dr. Philip Ngare and DR. Joseph Mwaniki for their guidance in writing this research paper. I also appreciate the support given to me by my classmates and friends.

REFERENCES

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

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimate of Variance of United Kingdom inflation. *Econometrica,* 40, 987-1007.

Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48(5):1779-1801.

Nelson, D. B. (1991). Conditional Heteroscedasticity in Asset returns: A New Approach, *Econometrica*, 59 , 347 -370.

Ngailo, E., Luvanda, E. and Massawe, E.S. (2014) Time Series Modelling with Application to Tanzania Inflation Data. Journal of Data Analysis and Information Processing, 2, 49-59. http://dx.doi.org/10.4236/jdaip.2014.22007

Webster, D. (2000). Webster's New Universal Unabridged Dictionary. Barnes & Noble Books, New York.