Modeling USD/KES Exchange Rate Volatility using GARCH Models

Cyprian Ondieki Omari[1*], Peter Nyamuhanga Mwita[2], Antony Gichuhi Waititu[2]

[1]*Department of Statistics and Actuarial Science, Dedan Kimathi University of Technology, Nyeri, Kenya*
[2]*Department of Statistics and Actuarial Sciences Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya*

Abstract: *In this paper the generalized autoregressive conditional heteroscedastic models are applied in modeling exchange rate volatility of the USD/KES exchange rate using daily observations over the period starting 3rd January 2003 to 31st December 2015. The paper applies both symmetric and asymmetric models that capture most of the stylized facts about exchange rate returns such as volatility clustering and leverage effect. The performance of the symmetric GARCH (1, 1) and GARCH-M models as well as the asymmetric EGARCH (1, 1), GJR-GARCH (1, 1) and APARCH (1, 1) models with different residual distributions are applied to data. The most adequate models for estimating volatility of the exchange rates are the asymmetric APARCH model, GJR-GARCH model and EGARCH model with Student's t-distribution.*
Keywords: *GARCH Models, Volatility clustering, forecasting volatility, Leverage effect, Value-at-Risk*

I. Introduction

In the last decade the foreign exchange market has become the most volatile and liquid in all financial markets in the world. Particularly because of the dynamics of the foreign exchange market, it is essential to study some of the important historical events relating to currencies and currency exchange. The modeling and forecasting exchange rates volatility has important implications in a range of areas in macroeconomics and finance. Value-at-Risk (VaR) is a risk measurement tool based on loss distributions. The Basel III framework developed by the Basel Committee on Banking Supervision requires that financial institutions such as banks and investment firms set aside a minimum amount of capital to cover potential losses from their exposure to credit risk, operational risk and market risk. For measuring market risk they recommend the use of VaR, which is the worst loss in an asset or a portfolio of assets over a given time horizon at a given confidence level. Inaccurate portfolio VaR estimates may lead firms to maintain insufficient risk capital reserves so that they have an inadequate capital cushion to absorb large financial shocks.

A number of models have been developed in empirical finance literature to investigate volatility across different regions and countries. The most commonly applied models to estimate exchange rate volatility are the autoregressive conditional heteroscedastic (ARCH) model introduced by Engle (1982) and the generalized (GARCH) models developed independently by Bollerslev (1986) and Taylor (1986). The purpose of the autoregressive conditional heteroscedasticity (ARCH) model is to estimate the conditional variance of a time series. Engle described the conditional variance by a simple quadratic function of its lagged values. Bollerslev (1986) extended the basic ARCH model and described the conditional variance by its own lagged values and the square of the lagged values of the innovations or shocks. In many cases, the basic GARCH model provides a reasonably good model for analyzing financial time series and estimating conditional volatility. However, GARCH models have been criticized in that they do not provide a theoretical explanation of volatility or what information flows are in the volatility generating process according to Tsay (2010). The GARCH model also responds equally to asymmetric shocks, and cannot cope with significantly skewed time series which results in biased estimates. Another problem encountered when using GARCH models is that they do not always fully embrace the heavy tails property of high frequency financial time series. To overcome this drawback Bollerslev et al. (1987) used the Student's *t* distributions. The GARCH extensions such as Exponential GARCH, Threshold GARCH, GJR-GARCH model and power GARCH models have been proposed to address some of these weaknesses. Nelson (1991) formulated the Exponential GARCH (EGARCH) model by extending the GARCH model to capture news in the form of leverage effects. Afterwards, the GARCH model extension was developed to test for this asymmetric news impact (Glosten et al., 1993; Zakoian, 1994). Ding et al. (1993) extensions nest a number of models from the ARCH family. The GARCH family models capture heteroscedasticity and volatility clustering in financial data.

The main objective of this paper is to model exchange rate return volatility for USD/KES, by applying different univariate specifications of GARCH type models for daily observations of the exchange rate return series for the period 3rd January 2003 to 31st December 2015. The volatility models applied in this paper include the GARCH (1, 1), GARCH-M (1, 1), E-GARCH (1, 1), GJR-GARCH (1, 1), and Power GARCH (1, 1). The

relative performance of the symmetric and asymmetric GARCH family models in estimating and forecasting Value-at-Risk using the USD/KES exchange rates is also tested.

The remainder of this paper is organized as follows. Section 2 provides the overview of symmetric and asymmetric GARCH family models used throughout the paper. Section 3 describes the data and empirical results and finally, Section 4 concludes the paper.

II. Methodology

The traditional methods of measuring volatility (variance or standard deviation) are unconditional and cannot capture the characteristics exhibited by financial time series data, such as, time varying volatility, volatility clustering, excess kurtosis, heavy tailed distribution, leverage effect and long memory properties. The most commonly used models that capture these properties of financial time series data are the Autoregressive Conditional Heteroskedasticity (ARCH) model and its generalization, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. In this paper different univariate GARCH specifications are applied to model USD/KES exchange rate return volatility and these models are GARCH (1, 1), GARCH-M (1, 1), EGARCH (1, 1), GJRGARCH (1, 1) and APARCH (1, 1). In presenting these different models, there are two distinct equations or specifications, the first the conditional mean and the conditional variance which are briefly reviewed in this methodology.

2.1 Conditional Mean Equation

The exchange rate return moving pattern might be an autoregressive (AR) process, moving average (MA) process or a combination of AR and MA processes i.e. (ARMA) process. For the purposes of this study the mean equation is modified to include appropriate AR and MA terms to control for autocorrelation in the data. For example, in ARMA (1, 1) process pattern would be:

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j}$$

(1)

where Y_t is a time series being modeled.

2.2 Volatility Modeling

The existing models of volatility can be divided into two main categories, symmetric and asymmetric models. In the symmetric models, the conditional variance only depends on the magnitude, and not the sign, of the underlying asset, while in the asymmetric models the positive or negative shocks of the same magnitude have different effect on future volatility.

2.3 Symmetric GARCH Models
2.3.1 Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Model

In this model, the conditional variance is represented as a linear function of its own lags. The general form of the GARCH *(p, q)* model is given by:

$$r_t = \mu + y_t, \qquad y_t = \sigma_t \varepsilon_t$$

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

(2)

where r_t are the logarithm returns of the financial time series at time *t*, μ are mean value of the returns, y_t are the error terms (innovations) from the mean equation, and it can split into a stochastic piece ε_t and a time dependant standard deviation σ_t characterizing the typical size of the terms. ε_t is a zero mean, identical and independent distribution, which is assumed to have normal distribution, *t* distribution and skew *t* distribution and $\omega > 0, \alpha_i \geq 0, \beta_j \geq 0,$ with constrains that $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$. In most empirical applications the basic GARCH (1, 1) model fits the changing conditional variance of the majority of financial time series reasonably well. The GARCH (1, 1) model is given by the following equation:

$$\sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

(3)

To guarantee a positive variance at all instances, the following restricts are imposed $\omega > 0$ and $\alpha_1, \beta_1 \geq 0.$

In many cases, the basic GARCH model provides a reasonably good model for analyzing financial time series and estimating conditional volatility. However, there are some aspects of the model which can be improved so that it can better capture the characteristics and dynamics of a particular financial time series. Ever since the GARCH *(p, q)* model was introduced by Bolleslev (1986), new GARCH extension models which address the different weaknesses of the GARCH model and capture different characteristics of the financial time series data have been invented.

2.3.2 The GARCH-in-Mean (GARCH-M) Model

In finance, high risk is often expected to lead to high returns. To model such a phenomenon one may consider the GARCH-M Model of Engle, Lilien, and Robins (1987) where "M" stands for GARCH in the mean. This model is an extension of the basic GARCH framework which allows the conditional mean of a sequence to depend on its conditional variance or standard deviation. A simple GARCH-M (1, 1) model is given by:

$$r_t = \mu + \lambda \sigma_t^2 + y_t, \qquad y_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim N\left(0, \sigma_t^2\right)$$
$$\sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{4}$$

where μ and λ are constants. The parameter λ is called the risk premium parameter. A positive λ indicates that the return is positively related to its volatility.

2.4 Asymmetric GARCH Models

In practice, the price of financial assets often reacts more pronouncedly to "bad" news than "good' news. Such a phenomenon leads to a so called leverage effect, as first noted by Black (1976). The term "leverage" stems from the empirical observation that the volatility (conditional variance) of a stock tends to increase when its returns are negative. The leverage effect causes the asymmetries of variance dynamics and points out the drawbacks of GARCH model because of its symmetric effect towards the conditional variance. In order to capture the asymmetry in return volatility ("leverage effect"), a new class of models was developed, termed the asymmetric GARCH models. This paper uses the following asymmetric GARCH models; EGARCH GJR-GARCH and Asymmetric Power ARCH (APARCH) model for capturing the asymmetric phenomena.

2.4.1 The Exponential GARCH (E-GARCH) Model

The general form of the Exponential GARCH *(p, q)* model introduced by Nelson (1991) is given by

$$\ln (\sigma_t^2) = \omega + \sum_{i=1}^{p} \alpha_i \left| \frac{y_{t-i}}{\sigma_{t-i}} \right| + \gamma_i \frac{y_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2 \tag{5}$$

where γ is the asymmetric response parameter that can take a positive or negative sign depending on the effect of the future uncertainty. The simplest form is the EGARCH (1, 1) model, which is given by:

$$\ln \sigma_t^2 = \omega + \alpha_1 \left| \frac{y_{t-1}}{\sigma_{t-1}} \right| + \gamma_1 \frac{y_{t-1}}{\sigma_{t-1}} + \beta_1 \ln \sigma_{t-1}^2 \tag{6}$$

For a positive shock $\dfrac{y_{t-1}}{\sigma_{t-1}} > 0,$ the equation becomes

$$\ln \sigma_t^2 = \omega + (\alpha_1 + \gamma_1) \frac{y_{t-1}}{\sigma_{t-1}} + \beta_1 \ln \sigma_{t-1}^2 \tag{7}$$

whereas for a negative shock $\dfrac{y_{t-1}}{\sigma_{t-1}} < 0,$ the equation becomes

$$\ln \sigma_t^2 = \omega + (\alpha_1 - \gamma_1) \frac{y_{t-1}}{\sigma_{t-1}} + \beta_1 \ln \sigma_{t-1}^2 \tag{8}$$

2.4.2 The Glosten, Jagannathan and Runkle GARCH (GJR-GARCH) Model

The GJR-GARCH model is another type of asymmetric GARCH models, which was proposed by Glosten, Jagannathan and Runkle (1993). The variance equation of GJR-GARCH *(p, q)* is given by

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \gamma_i I_{t-i} y_{t-i} \tag{9}$$

---

where α, β and γ are constant parameters, and I is a dummy variable (indicator function) that takes the value zero (respectively one) when y_{t-i} is positive (negative). If γ is positive, negative errors are leveraged (negative innovations or bad news has a greater impact than the positive ones). The parameters of the model are assumed to be positive and that $\alpha + \beta + \gamma/2 < 1$. If all leverage coefficients are zero, then GJR-GARCH model reduces to GARCH model. This means one can test a GARCH model against a GJR-GARCH model using the likelihood ratio test.

2.4.3 The Power GARCH (PGARCH) Model
Ding, Engle and Granger (1993) introduced the asymmetric power ARCH model also called APARCH *(p, q)* specification to deal with asymmetry. The variance equation of APARCH *(p, q)* can be written as

$$\sigma_t^{\delta} = \omega + \sum_{i=1}^{p} (\alpha_i |y_{t-i}| - \gamma_i y_{t-i})^{\delta} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^{\delta} \tag{10}$$

where $\omega > 0$, $\delta > 0$, $\alpha_i \geq 0$, $-1 < \gamma_i < 1$, $i = 1, \ldots, p$, $\beta_j \geq 0$, $j = 1, \ldots, q$. α_i and β_i are the standard ARCH and GARCH parameters, γ_i are the leverage parameters and δ is the parameter for the power term. The symmetric model sets $\gamma_i = 0$ for all *i*. When $\delta = 2$, Equation (10) becomes a classic GARCH model that allows for leverage effects and when $\delta = 1$, the conditional standard deviation will be estimated. In addition, we can increase the flexibility of the APARCH model by considering δ as another coefficient that must also be estimated.

In this paper, conditional volatility is estimated using the probability distributions that are available in the rugarch package which include; normal, Student *t* and skewed Student *t*-distribution. Engle (1982) assumed that asset returns follow a normal distribution. However, the asset returns are not normally distributed, so the normality assumption could cause significant bias in VaR estimation and could underestimate the volatility. A number of authors evidenced that standard GARCH models with normal empirical distributions have inferior forecasting performance compared to models that reflect skewness and kurtosis in innovations. To capture the excess kurtosis in financial asset returns, Bollerslev (1987) introduced the GARCH model with a standardized Student's *t* distribution with $\upsilon > 2$ degrees of freedom.

The common methodology used for GARCH estimation is maximum likelihood assuming i.i.d. innovations. The parameters of the GARCH model can be found by maximizing the objective log-likelihood function:

$$\ln L(\theta) = -\frac{1}{2} \sum_{t=1}^{n} \left[\ln(2\pi) + \ln(\sigma_t^2(\theta)) + z_t^2(\theta) \right] \tag{11}$$

where θ is the vector of parameters $(\omega, \mu, \alpha_i, \beta_j)$ estimated that maximize the objective function $\ln L(\theta)$; z_t represents the standardized residual calculated as $\dfrac{\Delta y_t - \mu}{\sqrt{\sigma_t^2}}$.

Maximum likelihood estimates of the parameters are obtained by numerical maximization of the log-likelihood function using the Marquardt algorithm (Marquardt (1963)). We use the quasi-maximum likelihood estimator (QMLE) since, according to Bollerslev and Wooldridge (1992), it is generally consistent, has a normal limiting distribution and provides asymptotic standard errors that are valid under non-normality.

For the GARCH *(p, q)* model the one-step-ahead conditional variance forecast, $\hat{\sigma}_{t+1|t}$ is:

$$\hat{\sigma}_{t+1|t}^2 = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i+1}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j+1}^2 \tag{12}$$

For the EGARCH *(p, q)* model, we get:

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \left| \frac{y_{t-i+1}}{\sigma_{t-i+1}} \right| + \gamma_i \frac{y_{t-i+1}}{\sigma_{t-i+1}} + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j+1}^2 \tag{13}$$

Note that the value of $E|z_t|$ depends on the density function of z_t. For example, for the standard normal distribution, $E|z_t| = \sqrt{2/\pi}$, for the Student *t*-distribution,

$$E \mid z_t \mid = \frac{2\Gamma\left[(1+v)/2\right]\sqrt{v-2}}{1+(v-1)\Gamma(v/2)\sqrt{\pi}},$$

However, the quantity $y_{t-i+1}\sigma^{-1}_{t-i+1}$, both with and without the absolute value operator, for $i \geq 1$, can be computed by the model as the values of both the innovation and its conditional standard deviation are available. Finally, the corresponding one-step ahead conditional variance forecast in the case of the GJR-GARCH (p, q) model is:

$$\sigma^2_{t+1|t} = \omega + \sum_{i=1}^{p}\alpha_i y_{t-i+1} + \sum_{j=1}^{q}\beta_j \sigma^2_{t-j+1} + \gamma_i I_{t-i+1} y_{t-i+1}$$
(14)

Therefore, to compute the one-step-ahead VaR forecast under all distributions, we compute the corresponding quantiles, which are then, multiply by the conditional standard deviation forecast, hence;

$$VaR_{t+1|t} = F(\alpha)\hat{\sigma}_{t+1|t},$$
(15)

where $F(\alpha)$ is the corresponding quantile of the assumed distribution, and $\hat{\sigma}_{t+1|t}$ is the forecast of conditional standard deviation at time t.

According to Tsay (2010), if one further assumes that z_t is Gaussian, then the conditional distribution of r_{t+1} given the information available at time t is $N\left(\hat{r}_{t+1|t}, \hat{\sigma}_{t+1|t}\right)$. Quantiles of this conditional distribution can easily be obtained for VaR calculation. For example, the 5% quantile is

$$\hat{r}_{t+1|t} - 1.645\hat{\sigma}_{t+1|t}.$$

Therefore, if z_t of the GARCH model in Equation (12) is a standardized Student-t distribution with v degrees of freedom and the probability is p, then the quantile used to calculate the 1-period horizon VaR at time index t is

$$\hat{r}_{t+1|t} - \frac{t_v(p)\hat{\sigma}_{t+1|t}}{\sqrt{v/(v-2)}}$$

where $t_v(p)$ is the p-th quantile of a Student-t distribution with v degrees of freedom.

III. Empirical Results

3.1 Data

The data set consists of the daily currency exchange rate of the US Dollar versus Kenyan Shilling (USD/KSH). These data are obtained from Central Bank of Kenya (CBK) website, (www.cbk.co.ke). The data set was for the period from January 3, 2003 to December 31, 2015, a total of 2818 observations. A visual inspection of Figure 1 shows that daily USD/KES exchange rate prices are not stationary. In order to test for stationarity an Augmented Dickey–Fuller test (ADF) for a unit root in a time series sample is performed. The computed ADF test-statistic in Table 1 is (-3.0) which greater than the critical values at one per cent significance level. Therefore, we fail to reject the null hypothesis that there is a unit root and that the series needs to be differenced in order to make it stationary.

Table 1: Augmented Dickey-Fuller test of the daily returns

ADF Test Statistic -3.0		
	1% Critical value	-3.4360
	5% Critical value	-2.8632
	10% Critical value	-2.5677

The currency exchange rates are then transformed into daily log returns using the following returns formula:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) \times 100$$
(16)

where P_t is the daily closing value of the USD/KES exchange rate on day *t*.

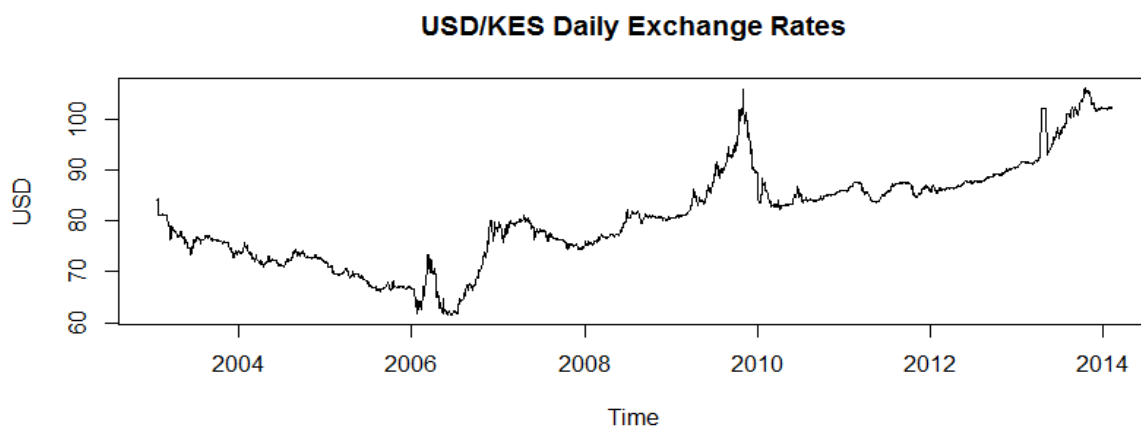**USD/KES Daily Exchange Rates**

Figure 1: Daily USD/KES Currency Exchange Rates

A plot of the log returns series for USD/KES exchange rates given in Figure 2 shows periods of high volatility, occasional extreme movements and volatility clustering, as upward movements tend to be followed by other upward movements and downward movements also followed by other downward movements. This indicates that the logarithm of USD/KES exchange rates is stationary after taking the first-difference, and the ADF test results in Table 2 confirm the stationarity of the return series data. The computed ADF test-statistic in Table 2 is (-10.0) which smaller than the critical values at 5% significance level.

Table 2: Augmented Dickey-Fuller test of the daily returns

ADF Test Statistic	1% Critical value	-3.4360
-10		
	5% Critical value	-2.8632
	10% Critical value	-2.5677

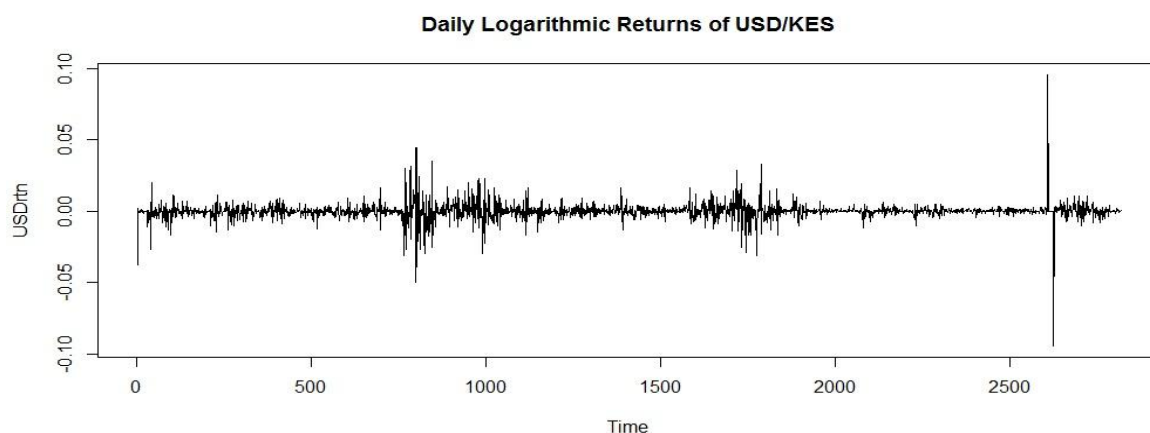**Daily Logarithmic Returns of USD/KES**

Figure 2: Daily Logarithmic Returns of the USD/KES Currency Exchange Rates

A summary of the statistics of the return series data is given in Table 3. The mean is positive, suggesting that exchange returns increase slightly over time. The coefficient of skewness indicates that returns have asymmetric distribution, i.e., they are skewed to the left. The kurtosis of returns is 73.6776 which is greater than three, indicating that the distribution of returns follows a fat-tailed distribution, thereby exhibiting one of the important characteristics of financial time series data, namely that of leptokurtosis. The non-normality

condition is supported by a Jarque-Bera test which shows that the null hypothesis of normality is rejected at the five per cent level of significance.

Table 3: Summary Statistics for the returns of USD/KES exchange rates

Number of Observations	2817
Max.	9.5776
Min.	-9.4458
Mean	0.0069
Median	0.0053
Variance	0.3054
Std.Dev	0.5527
Skewness	-0.0736
Kurtosis	73.6776
Jarque-Bera	638102.8932
JB p-value	*0.0000*

The Ljung-Box test is applied to the daily log returns of the USD/KES exchange rates and the test results are shown in Table 4. The null hypothesis of the Ljung-Box is rejected for the returns, squared returns and absolute returns, at lags 1, 6, 10, 15 and 20. The test statistics are statistically significant with p-values not greater than 0.01, indicating that the returns are not white noise. Indeed, the daily exchange rate returns exhibits correlation.

Table 4: p-values based on the Ljung-Box test for of the USD/KES exchange rates

	m	1	6	10	15	20
Returns	Q_m	10	40	50	60	90
	p-value	*(0.0004)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
Squared returns	Q_m	6	20	30	40	600
	p-value	*(0.01)*	*(0.0004)*	*(0.002)*	*0.003*	*(0.0000)*
Absolute returns	Q_m	200	800	1000	1000	2000
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*

From the results of Ljung-Box test in Table 4 and the autocorrelation (ACF) and partial autocorrelation (PACF) plots in Figure 3, for the exchange rate return series, absolute and squared return series shows that the return series exhibit autocorrelation at some lags at 5% level of significance. The presence of autocorrelation detected in the log return can be removed by fitting the simplest plausible ARMA *(p, q)* model to the data. On the other hand, the autocorrelation detected in the squared log returns, indicate that there exists conditional heteroskedasticity of the exchange rate returns series which could be removed by fitting the simplest plausible GARCH model to the ARMA filtered data.

3.2 Estimated Mean Equation

An ARMA *(p, q)* model is used to fit the mean returns, as it provides a flexible and parsimonious approximation to conditional mean dynamics. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) are used to determine the order of ARMA *(p, q)* models. The ACF and PACF plots given in Figure 3 suggest that the returns may be modeled by an ARMA (2, 2) process. Tsay and Tiao (1984) proposed the extended autocorrelation function (EACF) technique to identify the orders of a stationary or non-stationary ARMA process based on iterated least square estimates of the autoregressive parameters. The output of EACF is a two-way table, where the rows correspond to AR order *p* and the columns to MA order *q*. Therefore, the EACF suggests that the daily log returns of USD/KES exchange rate follow an ARMA *(2, 0)* model. This is in agreement with the result in Table 5 suggested by the best fitting model selected based on Bayesian Information Criterion (BIC) values. The criterion is to choose a model with minimum AIC and BIC and largest log-likelihood function. BIC always gives penalty for the additional parameters more than AIC does. So the ARMA (2, 0) is selected as the mean equation that mainly takes account of the BIC.

Table 5: Criterion for ARMA (p, q) Order Selection

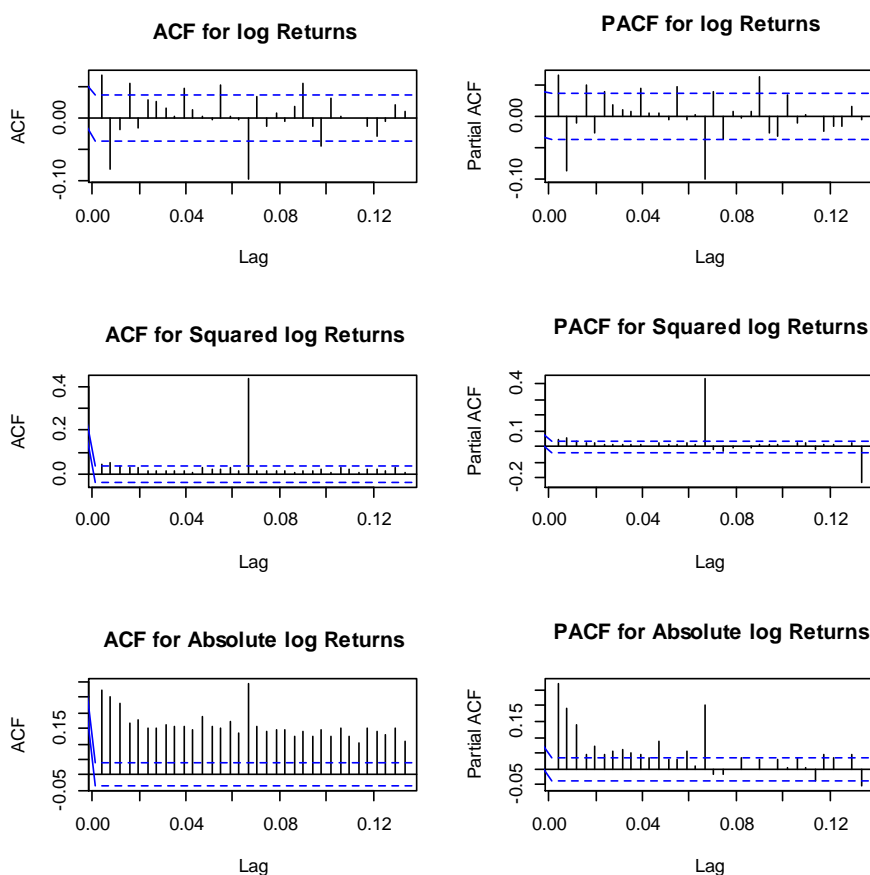|	AIC	BIC
ARMA (0, 0)	-21289	-21277
ARMA (1, 0)	-21300	-21282
ARMA (2, 0)	**-21319**	**-21296**
ARMA (3, 0)	-21318	-21288
ARMA (0, 1)	-21303	-21285
ARMA (0, 2)	-21316	-21292
ARMA (0, 3)	-21317	-21287
ARMA (1, 1)	-21307	-21283
ARMA (1, 2)	-21315	-21285
ARMA (2, 1)	-21317	-21288
ARMA (2, 2)	-21324	-21288

Figure 3: ACF/PACF Plots for USD/KES returns

3.3 Estimated Volatility Model

The results of the fitted AR (2)-GARCH (1, 1) and AR (2)-GARCH –M (1, 1) models to the USD/KES log return series with normal distribution, Student's t distribution and skewed t distribution for the standardized residuals are presented in Table 6. The estimates of the model parameters are all significant for normal, Student's t and skewed t distribution except for the ω parameter which is not significant for all the distributions.

The estimates of ϕ_1 and ϕ_2 are significant, supporting the use of the AR (2) model for the returns. Volatility shocks are persistent since the sum of the ARCH and GARCH coefficients are very close to one. The Box-Pierce Q statistics is insignificant up to lag 20, indicating that there is no excessive autocorrelation left in the residuals. Comparing the log-likelihood and information criterion in Table 6 within the three conditional distributions, the model with conditional distribution of skewed t has larger log-likelihood and smaller information criterion statistics than estimated by normal and t distribution which means this model is better fitted.

Table 6: Estimation of AR (2)-GARCH (1, 1) and AR (2)-GARCH-M (1, 1) with Different Distributions

	AR (2)-GARCH (1, 1)			AR (2)-GARCH-M (1, 1)		
	Normal	**t**	**Skew t**	**Normal**	**t**	**Skew t**
μ	0.000137	0.000079	0.000093	0.000141	0.000082	0.000094
	(0.00230)	*(0.01642)*	*(0.017355)*	*(0.005763)*	*(0.018374)*	*(0.017024)*
AR(1)	0.132993	0.139786	0.139081	0.131115	0.139499	0.138878
	(0.00000)	*(0.00000)*	*(0.00000)*	*(0.00000)*	*(0.00000)*	*(0.00000)*
AR(2)	-0.097567	-0.044787	-0.045087	-0.098027	-0.044807	-0.045080
	(0.00000)	*(0.012145)*	*(0.011554)*	*(0.000010)*	*(0.012114)*	*(0.011577)*
Omega	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	(0.658234)	*(0.469487)*	*(0.472896)*	*(0.675454)*	*(0.470185)*	*(0.473492)*
α_1	0.120459	0.318323	0.317466	0.118585	0.318264	0.317520
	(0.00000)	*(0.00000)*	*(0.00000)*	*(0.00000)*	*(0.00000)*	*(0.00000)*
β_1	0.878420	0.680677	0.681534	0.880377	0.680736	0.681480
	(0.00000)	*(0.00000)*	*(0.00000)*	*(0.00000)*	*(0.00000)*	*(0.00000)*
Skew						1.011980
						(0.00000)
Shape		3.146979	1.013131		3.147627	3.150416
		(0.00000)	*(0.00000)*		*(0.00000)*	*(0.00000)*
LLF	12002	12511	12512	12002	12511	12512
AIC	-8.5166	-8.8778	-8.8772	-8.5167	-8.8778	-8.8772
BIC	-8.5039	-8.8630	-8.8603	-8.5167	-8.8630	-8.8603

*P-values are shown in parentheses.

To capture the asymmetry dynamics and the presence of the "leverage effect" in the USD/KES exchange rate returns, the nonlinear asymmetric models; AR (2)-EGARCH (1, 1), AR (2)-GJR-GARCH (1, 1) and AR (2)-APARCH (1, 1) with conditional distributions; normal distribution, Student's t distribution and skewed t distribution are fitted to the exchange returns. Table 7 gives the results of the parameter estimates for the AR (2)-EGARCH (1, 1), AR (2)-GJR-GARCH (1, 1) and AR (2)-APARCH (1, 1) models. The parameters estimates for these three models are all significant except for the mean under the AR (2)-EGARCH (1, 1) for the normal and skew t distribution, also the coefficient of the second term of autoregressive process under the skew t distribution and the coefficients of α_1 under the Student's t and skew t distribution are not significant. For both the AR (2)-GJR-GARCH (1, 1) and AR (2)-APARCH (1, 1) ω is not significant for all the distribution. The parameter γ is not significant for the AR (2)-APARCH (1, 1) under the t distribution. The coefficient γ in the case of AR (2)-APARCH (1, 1) is statistically significant at level of significance of 5% implying that there is an asymmetry under the normal distribution. On the other hand, its negative value indicates the presence of the "leverage effect". The coefficient γ in the AR (2)-E-GARCH (1, 1) and AR (2)-GJR-GARCH (1, 1) is significantly different from zero, which indicates the presence of asymmetry. The value of γ which is less than zero implies presence of the "leverage effect". According to the log-likelihood value and information criterion of the estimated models, the APARCH model has the larger log-likelihood value and smaller information criterion compared with E-GARCH model and GJR-GARCH model. Secondly, comparing within the APARCH models under normal distribution, and Student's t distribution, the model with conditional Student's t distribution outperforms the normal distribution which means this model is superior in modeling the USD/KES exchange rate returns with asymmetry and fat tail.

The estimated power parameter δ in the APARCH model is 2.44 which is slightly different from the estimated result of Ding, Granger and Engle (1993)'s under the normal distribution which is 1.43. This may be caused by the time period of the data is different and then mean equation is also different to model the data. But δ in this paper is still significantly different from 1 (GJR-GARCH) or 2 (GARCH). When the conditional distribution changes to t distribution δ is getting smaller to 0.73, however, using the same test as in Ding, Granger and Engle (1993)'s paper, let l_0 be the log-likelihood of value under the GARCH model which is set as the null hypothesis, while the alternative hypothesis is APARCH model with log-likelihood is l, then $2(l - l_0)$ have a χ_2 distribution with 2 degrees of freedom when H_0 is true. Then, under the Student's t distribution $2(l - l_0) = 2(12547 - 12511) = 72$, which means we can reject the null hypothesis that the data is generated from GARCH model. And also in the same way we can reject that the data is generated from E-GARCH model and GJR-GARCH model.

Table 7: Estimation of AR (2)-EGARCH (1, 1) and AR (2)-GJR-GARCH (1, 1) with Different Distributions

	AR (2)-EGARCH (1, 1)			AR (2)-GJR-GARCH (1, 1)			AR (2)-APARCH (1, 1)	
	Normal	t	Skew t	Normal	t	Skew t	Normal	t
μ	-0.000138 (0.103591)	0.000077 (0.010444)	0.000085 (0.052036)	0.00154 (0.000768)	0.000084 (0.011170)	0.000100 (0.010824)	0.000100 (0.031989)	0.00059 (0.00000)
AR(1)	0.121891 (0.000036)	0.128796 (0.00000)	0.128516 (0.00000)	0.134449 (0.00000)	0.139695 (0.00000)	0.139224 (0.00000)	0.134804 (0.00000)	0.091779 (0.00000)
AR(2)	-0.046525 (0.000491)	-0.029849 (0.038131)	-0.029962 (0.076620)	-0.097030 (0.000011)	-0.043918 (0.013838)	-0.043998 (0.013590)	-0.100342 (0.000004)	-0.011551 (0.00000)
Omega	-0.535835 (0.00000)	-0.599479 (0.000097)	-0.599915 (0.000098)	0.000000 (0.665489)	0.000000 (0.482739)	0.000000 (0.489444)	0.000000 (0.981524)	0.000000 (0.962086)
α_1	0.111569 (0.00000)	0.022818 (0.453770)	0.022603 (0.452603)	0.13854 (0.00000)	0.354876 (0.00000)	0.354889 (0.00000)	0.087753 (0.00000)	0.394055 (0.00000)
β_1	0.942070 (0.00000)	0.944689 (0.00000)	0.944655 (0.00000)	0.879810 (0.00000)	0.686124 (0.00000)	0.687931 (0.00000)	0.879502 (0.00000)	0.855587 (0.00000)
Gamma	0.484044 (0.00000)	0.554240 (0.00000)	0.554074 (0.00000)	-0.025328 (0.026778)	-0.084006 (0.049376)	-0.086881 (0.041340)	-0.051172 (0.035613)	-0.085833 (0.066427)
Delta							2.437594 (0.00000)	0.727418 (0.00000)
Skew			1.006334 (0.00000)			1.014788 (0.00000)		
Shape		2.386515 (0.00000)	2.387376 (0.00000)		3.145621 (0.00000)	3.148859 (0.00000)		2.100008 (0.00000)
LLF	11557	12505	12505	12004	12513	12514	11997	12547
AIC	-8.2004	-8.8728	-8.8721	-8.5177	-8.8785	-8.8780	-8.5119	-8.9014
BIC	-8.1857	-8.8559	-8.8531	-8.5029	-8.8616	-8.8590	-8.4950	-8.8825

*P-values are shown in parentheses.

The GARCH models with the innovations of Student's *t* and skewed Student's *t* distributions have a better fit in general than the models with normal distribution innovations since they have the highest log-likelihood function (LLF) and smallest AIC and BIC. Secondly, the values of the AIC, BIC and LLF for all the models with Student's *t* and skewed Student's *t* distributions innovations are not significantly different. This implies that the models with Student's *t* and skewed Student's *t* distributions innovations would result in the same conclusions. The volatility (conditional variance) process and standardized residuals for the AR (2) - APARCH (1, 1) model with Student *t* distribution is plotted in Figure 4.
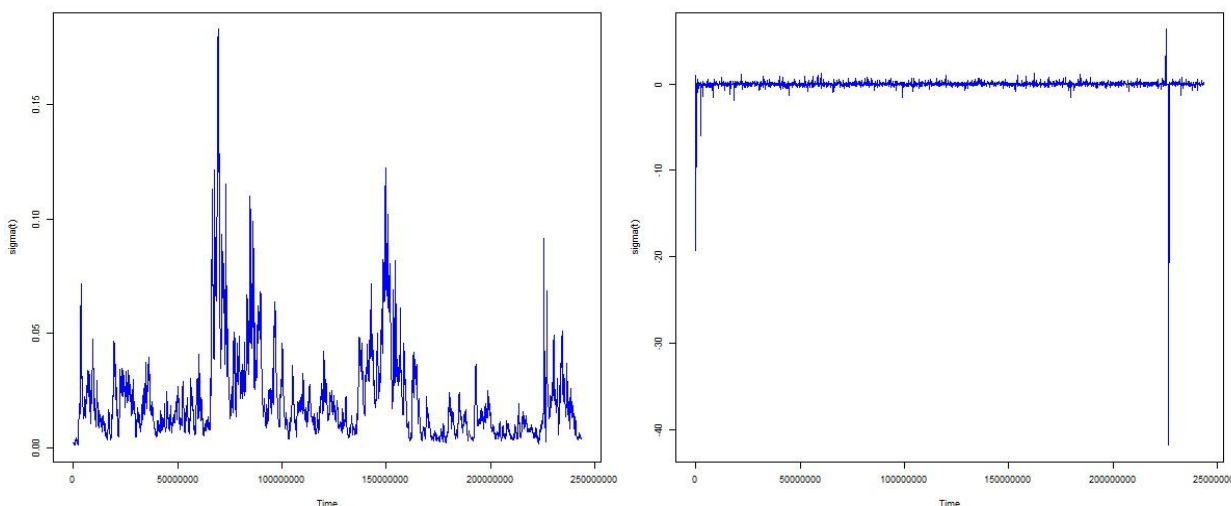

Figure 4: Volatility (conditional variance) process and standardized residuals of exchange rate returns derived from the AR (2)-APARCH (1, 1) residuals.

The plot in Figure 4 shows that the model is well specified. The ACF of the square standardized residuals compares well with the ACF of the square returns in Figure 3. This shows that AR (2)-APARCH (1, 1) Student-*t* model sufficiently explains the heteroscedasticity effect in the returns, thus we can conclude that the model fit the USD/KES returns well. The Ljung-Box test of the standardized residuals at different lags confirms that standardized residuals have no correlation.

The future return rate and volatility for one-day-ahead based on the estimated parameters of the models are obtained. These forecasted values are necessary for the estimation of Value at Risk (VaR). The estimated values of the VaR parameters for one-day-ahead as well as the probabilities of 95% and 99% are exhibited in Table 8.

Table 8: Econometric Estimation of the parameters of VaR for One-day-ahead period

Model	GARCH t distribution	GARCH-M t distribution	GJR-GARCH t distribution	EGARCH t distribution	APARCH t Distribution
Forecasted return	0.00005066	0.0000522	0.000055	0.00004574	0.00002786
Forecasted conditional variance	0.00127	0.00127	0.001269	0.00159	0.002049
VaR 0.95	-0.16%	-0.16%	-0.16%	-0.16%	-0.16%
VaR 0.99	-0.33%	-0.33%	-0.33%	-0.35%	-0.36%

The estimated VaR values obtained with the GARCH approach are negative. The negative sign is usually ignored since it's an indicator of loss. With probability of 0.95 the expected maximum loss due to having to change 1 US Dollars to KES is around in one day period.

IV. Conclusions

Modeling and forecasting the volatility of exchange rate returns has become an important field of empirical research in finance. This is because volatility is considered as an important concept in many economic and financial applications like asset pricing, risk management and portfolio allocation. This paper attempts to explore the comparative performance of different econometric volatility forecasting models in the terms of their ability to estimate VaR in the USD/KES exchange rates. A total of five different models were considered in this study. The volatility of the USD/KES returns have been modeled by using a univariate Generalized Autoregressive Conditional Heteroscedastic (GARCH) models including both symmetric and asymmetric models that captures most common stylized facts about exchange returns such as volatility clustering and leverage effect, these models are GARCH (1, 1), GARCH-M (1, 1), exponential GARCH (1, 1), GJR GARCH (1, 1) and APARCH (1, 1) following three residual distributions namely; normal, Student's *t*-distribution and Skewed Student's *t*-distribution. The first two models are used for capturing the symmetry effect whereas the second group of models is for capturing the asymmetric effect. The study used the USD/KES exchange rates data from the Central Bank of Kenya (CBK) for the period 3rd January 2003 to 31st December, 2015. Based on the empirical results presented, the following can be concluded:

The paper finds strong evidence that daily returns could be characterized by the above mentioned models. The USD/KES data showed a significant departure from normality and existence of conditional heteroscedasticity in the residuals series. Descriptive statistics for the USD/KES exchange rates show presence of negative skewness and excess kurtosis. The results of the conducted ARCH-LM test point out significant presence of ARCH effect in the residuals as well as volatility clustering effect. Standardized residuals and standardized residuals squared were white noise. The econometric estimation of VaR can be related to the chosen GARCH model. Therefore a first step in estimation of the parameters of VaR is a detailed specification analysis of the potential models. Based on the estimated model, a 1-step-ahead forecasting is taken to forecast the future value of the exchange rate returns and the conditional volatility. The values are used to estimate VaR. The empirical results have indicated that the most adequate GARCH models for estimating and forecasting VaR in the USD/KES exchange rates are the asymmetric APARCH, GJR-GARCH and EGARCH model with Student's *t*-distribution. These models have a better fit of the exchange returns, since they have the largest log-likelihood function and smallest AIC and BIC. We also compared the one step-ahead VaR estimate from the asymmetric models with Student's *t*-distribution and from the results the conclusion was that the AR (2)-APARCH (1, 1) model is also superior in the estimating the one-step-ahead VaR.

The findings in this paper have important implications regarding VaR estimation in volatile times, market timing, portfolio selection etc. that have to be addressed by investors and other risk managers operating in emerging markets. However, the limitation of the study is that the empirical research focused only on the USD/KES exchange rate and therefore the findings cannot be generalized to other exchange rates in the market. In the future research a wider sample of exchange rates should be used to compare the performance of the most commonly used foreign currencies in the market and the inclusion of other asymmetric GARCH-type models, testing and comparing their predictive performance.

References

[1] Black, F. (1976). Studies of stock price volatility changes.
[2] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307-327.
[3] Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. The review of *economics and statistics*, 542-547.
[4] Brooks, C. (2008). Introductory financial econometrics.

[5]   Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance*, *1*(1), 83-106.
[6]   Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.
[7]   Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica: Journal of the Econometric Society*, 391-407.
[8]   Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779-1801.
[9]   Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.
[10]   Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of econometrics*, *45*(1), 267-290.
[11]   Poon, S. H., & Granger, C. W. (2003). Forecasting volatility in financial markets: A review. *Journal of economic literature*, *41*(2), 478-539.
[12]   Tsay, R. S. (2010). *Multivariate time series analysis: with R and financial applications*. John Wiley & Sons.
[13]   Tsay, R. and Tiao, G. (1984). "Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Non-stationary ARMA Models." Journal of the American Statistical Association, 79 (385), pp. 84-96.
[14]   Taylor, S. (1986), Modelling Financial Time Series, John Wiley & Sons, Chichester
[15]   Zakoian, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, *18*(5), 931-955.