MODELLING THE VOLATILITY OF EXCHANGE RATES IN RWANDESE MARKETS

Jean de Dieu Ntawihebasenga[1], P.N.Mwita[2], J.K.Mung'atu[3]
[1]Pan African University Institute of Basic Sciences, Technology and innovation, Mathematics (Finance) Kenya,
[2]Jomo Kenyatta University of Agriculture and Technology, Department of Statistics and Actuarial Science, Kenya,
[3]Jomo Kenyatta University of Agriculture and Technology, Department of Statistics and Actuarial Science, Kenya,

ABSTRACT: *This work applied Generalized Autoregressive Conditional Heteroskedasticity (GARCH) approach to modelling volatility in Rwanda Exchange rate returns. The Autoregressive (AR) model with GARCH errors was fitted to the daily exchange rate returns using Quasi-Maximum Likelihood Estimation (Q-MLE) method to get the current volatility. Asymptotic consistency and asymptotic normality of estimated parameters were given. Akaike Information criterion was used for appropriate GARCH model selection while Jarque Bera test used for normality testing revealed that both returns and residuals have fat tails behaviour. It was shown that the estimated model fits Rwanda exchange rate returns data well.*

KEYWORDS: Model, Volatility, Exchange rate, Quasi Maximum Likelihood, GARCH model.

INTRODUCTION

Exchange rates are a challenging concept due to the fact that one has to deal with foreign exchange rates whenever he/she travels to a foreign country. Foreign exchange rates markets are world decentralized marketplaces that determine the relative values of different currencies. Volatility refers to the spread of all likely outcomes of an uncertain variable. Statistically, volatility is often measured as the sample standard deviation. Volatility is related to but not exactly the same as risk. Risk is related with undesirable results, but volatility can be defined as a measurement of the change in price over a given period of time. The conditional volatility in exchange rate returns is considered as the origin of exchange rates risk and has certain significances on the volume of international trade. Exchange rate volatility is a measure of the fluctuations in exchange rate markets.

The risk in foreign exchange can be defined as exposure to uncertainty and it cannot be dismissed in exchange rate markets since both importers and exporters of goods and services are affected by exchange rates fluctuations. Exchange rate risk also known as Foreign exchange risk or currency risk refers to a financial risk posed by an exposure to unanticipated changes in the exchange rate between two currencies. It may also be defined as the variability of a portfolio's value caused by uncertain fluctuations in the exchange rates. A value of any currency fluctuates as its demand and supply fluctuates, this means that if demand decrease or supply increase this can cause depreciation of the currency's value. On other hand if supply decreases and demand increases, this can cause

23

appreciation of the value of currency (Madura, 1989). When volatility in exchange rates increases it leads to uncertainty in pricing and this hurts importers who spend more for the same quantity. The volatility in prices has implications on the profits and survival of business enterprises Smith et al.(1990).

Recently, researchers and academics have estimated conditional volatility of exchange rates return using different techniques. Maana et al (2010) modeled exchange rate volatility of Kenya exchange rate markets using GARCH (1, 1) model. They found that the importers and exporters of goods and services are both affected by exchange rate fluctuations. Ghysels and Jasiak (1996) estimated the volatility as non-constant and non-symmetric with left fat tail. Some researchers argue that the true volatility cannot be estimated because there is no relationship between prior, current, and future volatilities for financial assets. If so, approaches utilizing volatility in estimating VaR should invalid, Sandmann & Koopman(1998).

Blum &Dacorogna (2002) studied extreme moves in daily Foreign Exchange rates and risk limit setting and they found that foreign exchange rates can be subject to considerable daily fluctuations (up to 5 percent within one day) and can cause serious losses on open overnight positions. They quantified the risk by concentrating their attention on the tails of the distribution and showed how to use estimations to compute limits that a risk manager can set to open positions to avoid unexpected huge losses.

Researchers in the past years have found out two fundamental attributes regarding distributions of returns of financial assets returns. First, most of the times, the distributions of financial assets returns are not normal, for example see, Hull & White (1998). Manganelli& Engle (2001) showed that financial returns, especially exchange rate and interest rate returns are not normally distributed, suffer from volatility clustering and are not independent. Some researchers argue that the distribution should have fat tails, Longin (1996) and Neftci (2000), and others argue that it should not be symmetric, Glosten (1993). Secondly, some researchers found that the distributions of financial assets returns are not constant over time. Such findings are related to another field of research in finance: the prediction of volatility of financial assets. There have been a lot of debate about the attributes of volatility; whether volatility is time-varying or constant, whether it should be weighed through time or not, or what time interval from the past is relevant for current volatility, Nelson (1991).

MATERIAL AND METHODS

This work applied GARCH approach to estimate exchange rate volatility of Rwanda Francs versus Japan Yen (Frw/JPY), Rwanda Francs against US Dollars (Frw/USD) and Rwanda Francs versus Sterling Pound (Frw/GBP) for the time period of 5 years. The daily data were obtained from The National Bank of Rwanda while R-software was used to generate the data. The choice of the currencies was based on their relative proportions in the Bank's foreign exchange investment portfolio and based also on its currency composition of imports. The data used were the average

of buying and selling of the Rwanda Exchange Rate covering the period from June 1st 2009 to June 1st 2014.

Volatility Estimation

In literature, different models in econometrics have been proposed to deal with pattern facts about volatility. These models include models from the ARCH/GARCH family. Engle (1982) and Bollerslev (1986) suggested ARCH and GARCH models respectively to resolve the problem of volatility clustering and persistence in financial data. ARCH/GARCH models play a crucial role in modeling of conditional volatility since they manage changing volatility with the assumption of conditional normality. Consider a sequence $\{e_t\}$ of independent and identically distributed random variable such that $e_t \sim N(0,1)$. ϵ_t is a GARCH (q, p) process if

$$\begin{cases} \epsilon_t = \sigma_t e_t, \epsilon_t \sim (0, \sigma_t^2) \\ \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \end{cases} \qquad (1)$$

The size of the parameters α_i and β_j determine the short run dynamics of the resulting volatility process. The non-negativity of α_i and β_j ensure that σ_t^2 is strictly positive. Large ARCH error coefficients α_i imply that volatility reacts significantly to market movements. Large GARCH coefficients β_j indicate that shocks to the conditional variance take long time to die out. High α_i coefficients, relative to β_j indicate that volatility tends to be more extreme. Since σ_t^2 is the one-period ahead forecast volatility based on the past information, it is called conditional volatility and it is specified as a function of three terms: unconditional volatility ω, news about volatility from the previous period measured as the lag of the squared residuals from the mean equation r_{t-1}^2 (ARCH term) and last period volatility σ_{t-1}^2 (GARCH term). The Akaike information criterion has been used to select a good GARCH model for each currency, Jarque and Bera (1982) came up with test for normality testing, this test has also been applied to see whether our data is normally distributed or not.. To test whether our data is stationary, this study has applied Augmented Dickey Fuller Test and Lagrange Multiplier test has been used for Arch effect testing in residuals. The GARCH (p, q) model was fitted to the daily exchange rate returns using Quasi-Maximum likelihood Estimate to get the current volatility.

$$\hat{\sigma}_t = \sqrt{\hat{\omega} + \sum_{i=1}^q \hat{\alpha}_i \epsilon_{t-i}^2 + \sum_{j=1}^p \hat{\beta}_j \sigma_{t-j}^2} \qquad (2)$$

Quasi-Maximum Likelihood Procedure

To get the Quasi-Likelihood function, we consider the situation where the true probability distribution $f_0(r_t, \theta_0)$ of the exchange rate at time t and incorrect probability distribution given by $f(r_t, \theta)$ are used to build the likelihood function. Now model (1) can be reformulated by letting $\{r_t\}$ be a sequence with the true model yield

$$\begin{cases} r_t = \epsilon_{0t}, \ \epsilon_{0t} = \sigma_{0t} e_t \\ \sigma_{0t}^2 = Var(r_t | \mathcal{F}_{t-1}) = E[(r_t)^2 | \mathcal{F}_{t-1}] \end{cases} \qquad (3)$$

Where $\epsilon_{0t} \sim N(0, \sigma_{0t}^2)$, $E(\epsilon_{0t} | \mathcal{F}_{t-1}) = 0$ *almost sure* $(a.s)$ and $\mathcal{F}_t = \sigma(\epsilon_{0t}, \epsilon_{0t-1}, \epsilon_{0t-2}, \dots)$, the conditional variance can be defined as $E(\epsilon_{0t}^2 | \mathcal{F}_{t-1}) = \sigma_{0t}^2$ (the subscript 0 indicates the true values of parameters). We also assume $r_t = \epsilon_t = \sigma_t \varepsilon_t$, $\epsilon_t \sim N(0, \sigma_t^2)$ to be the model for the unknown

parameters (misspecified model). Hence the true and misspecified distributions are;

$$f_0(r_t) = \frac{1}{\sigma_{0t}\sqrt{2\pi}} exp\left[-\frac{(\epsilon_{0t})^2}{2\sigma_{0t}^2}\right] \qquad (4)$$

$$f(r_t) = \frac{1}{\sigma_t\sqrt{2\pi}} exp\left[-\frac{(\epsilon_t)^2}{2\sigma_t^2}\right] \qquad (5)$$

Let us assume that innovations follow a GARCH (1, 1) process;

$$\sigma_{0t}^2 = \omega_0(1 - \beta_0) + \alpha_0\epsilon_{0t-1}^2 + \beta_0\sigma_{0t-1}^2 \; a.s$$

An equivalent expression for the conditional variance can be derived as:

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

we also assume that the process is described with true parameters in the vector form given as

$$\theta_0 = [\omega_0, \alpha_0, \beta_0]' \qquad (6)$$

and for the model with the unknown parameters,

$$\sigma_t^2(\theta) = \omega(1 - \beta) + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \quad t = 2,3,...,T$$

with the setup or initial condition $\sigma_1^2(\theta) = \omega$, this gives the convenient expression for the conditional variance process

$$\sigma_t^2 = \omega + \sum_{k=0}^{t-2} \beta^k \epsilon_{t-1-k}^2$$

Let us also assume that the innovation ϵ_t is the model for the unknown parameters

$$\theta = [\omega, \alpha, \beta]', \text{ with} |\beta| < 1 \qquad (7)$$

Now we can define the compact parameter space Θ, in the following way.

$$\Theta \equiv \{\theta: 0 < \omega_l \leq \omega \leq \omega_u; \; 0 < \alpha_l \leq \alpha \leq \alpha_u; \; 0 < \beta_l \leq \beta \leq \beta_u < 1\} \qquad (8)$$

Where subscript l and u indicate lower and upper limits respectively. We assume that the true parameter $\theta_0 \in \Theta$, this implies that $\alpha_0 > 0$, $\beta_0 > 0$, which means that ϵ_t is strictly a GARCH process. We can also define standardized residuals $e_t = {\epsilon_t}/{\sigma_t}$ by constructing $E(e_t|\mathcal{F}_{t-1}) = 0 \; a.s$ and $E(e_t^2|\mathcal{F}_{t-1}) = 1 \; a.s$ frequently, estimation of GARCH models is done under the assumption that $e_t \sim i.i.d \; N(0,1)$ so that the likelihood is easily specified. The maximum likelihood estimators of the parameters of the misspecified distribution are obtained by maximizing the log-likelihood function

$$lnL(\theta) = \sum_{t=1}^n lnf(r_t; \theta) \qquad (9)$$

The $\hat{\theta}$ estimator is obtained by setting the first order conditions given by

$$l(\theta) = \frac{\partial lnL}{\partial\theta} = \sum_{t=1}^n \frac{\partial lnf(r_t; \theta)}{\partial\theta} \qquad (10)$$

to zero. Let us take expectations of the gradient vector in (10) with respect to the true probability distribution $f_0(r_t; \theta_0)$.

$$E_0[l(\theta)] = \int_{-\infty}^{\infty} l(\theta) f_0(r_t; \theta_0)dr_t$$

$$= \int_{-\infty}^{\infty} \sum_{t=1}^n \frac{\partial lnf(r_t; \theta)}{\partial\theta} f_0(r_t; \theta_0)dr_t$$

$$= \int_{-\infty}^{\infty} \sum_{t=1}^n \frac{\partial f(r_t; \theta)}{\partial\theta} \frac{f_0(r_t; \theta_0)dr_t}{f(r_t; \theta)} \qquad (11)$$

Where $E_0[.]$ means that the expectation taken with respect to the true distribution, this expression is not guaranteed to equal zero except in the case where the distribution is specified correctly $f(r_t; \theta) = f_0(r_t; \theta_0)$. In this case, (11) may be simplified by exchanging the integration and differentiation operators and using the property of a probability distribution to give

$$E_0[l(\theta)] = \sum_{t=1}^n \frac{\partial}{\partial\theta} \int_{-\infty}^{\infty} f(r_t; \theta) \, dr_t$$

$$= \sum_{t=1}^n \frac{\partial}{\partial\theta} 1$$

$$E_0[l(\theta)] = 0 \qquad (12)$$

Thus sufficient condition for (12) to hold is that the model is specified correctly. There are however, some important cases where $E_0[l(\theta)] = 0$ even when the distribution is misspecified. Let us assume that the Gaussian Likelihood is applied to form the estimator. Then, the Log Likelihood takes the form

$$L_n(\theta) = \frac{1}{2n}\sum_{t=1}^{n} l_t(\theta), \qquad (13)$$

Where $l_t(\theta) = -\left[\ln \sigma_t^2(\theta) + \frac{\epsilon_t^2}{\sigma_t^2(\theta)}\right]$ and $L_n(\theta)$ is typically referred to as a quasi-likelihood function of parameter θ, since the likelihood need not to be the correct density. The vector of parameter value, denoted by $\hat{\theta}_n$ maximizes the likelihood Equation (13) on the subspace Θ_I belongs to compact space Θ ($\Theta_I \subset \Theta$) is obtained as:

$$\hat{\theta}_n = arg \max_{\theta \in \Theta_I} L_n(\theta) \qquad (14)$$

We also need to investigate the asymptotic consistency and asymptotic normality properties of the quasi-maximum estimator $\hat{\theta}_n$ of the GARCH process. To study the asymptotic properties of our estimators, we may discuss two important questions in asymptotic properties of our estimators: first, does our estimator work properly asymptotically? i.e, does it provide as with the correct answer if we give it enough data? The second, how asymptotic efficient is our estimator?

Asymptotic Consistency
An estimator say $\hat{\theta}_n$ is consistency to the actual parameter θ_n means that when sample size is sufficiently large the estimator $\hat{\theta}_n$ will be very likely to be very close to the actual parameter value θ_n. When an estimator converges in probability to the true value as the sample size increases, then, the estimator is asymptotically consistent. In this study, let us suppose that we observe the daily exchange rate returns data $r_{-p+1}, \dots r_0, r_1, \dots, r_n$ generated by the model (3) with θ_0 as the parameter. Assume that the data up to r_0 are available to us and the process r_0 is described with true parameters in the vector form given as in (6). One can define

$$\hat{\sigma}_{*t}^2(\theta) = \omega(1-\beta) + \alpha\hat{\epsilon}_{t-1}^2 + \beta\hat{\sigma}_{*t-1}^2 \quad t = 1,2,..T \qquad (15)$$

together with initialization $\hat{\sigma}_{0t}^2(\theta) \geq 0$ this means that the log-likelihood of $(r_1, \dots, r_n)'$ given $(r_0, \sigma_0)'$ under $e_t i.i.d \sim N(0,1)$ is approximately equal to

$$\hat{L}_n(\theta) = -\frac{1}{2}\sum_{t=1}^{n}\left(\ln\hat{\sigma}_{*t}^2(\theta) + \frac{r_t^2}{\hat{\sigma}_{*t}^2(\theta)}\right) \qquad (16)$$

The Quasi-Maximum Likelihood Estimator $\hat{\theta}_n$ is the parameter value which maximizes $\hat{L}_n$ on subspace Θ_1, since Θ_1 is an approximately chosen compact subset of the parameter space Θ. Therefore, Quasi-MLE $\hat{\theta}_n$ is strongly consistency if the following conditions on the random variable e_t are satisfied:

D_1. e_t is sequence of independent and identically distributed random variables such that E(e_t)=0
D_2. The vector parameter θ_0 is in the interior of compact set Θ.
D_3. For some a>0 there exists a constant $b < \infty$ such that $E[e_t^{2+a}] \leq b < \infty$
D_4. e_t^{2+a} is non degenerate
D_5. $E[ln(\beta_0 + \alpha_0 e_t^2)] < 0$
D_6. If for some t holds $\sigma_{0t}^2 = \omega_0 + \sum_{k=1}^{\infty} \omega_k \epsilon_{t-k}^2$ and $\sigma_{0t}^2 = \omega_0^* + \sum_{k=1}^{\infty} \omega_k^* \epsilon_{t-k}^2$
Therefore $\omega_j = \omega_j^*$ for every $1 \leq j < \infty$

If these conditions are satisfied we can conclude the consistency of Quasi-MLE in the following theorem.

Theorem

Under the conditions $D_1.-D_6$ above, the quasi-maximum likelihood estimate $\hat{\theta}_n$ is strongly consistent that is

$$\hat{\theta}_n \xrightarrow{a.s} \theta_0, \; n \to \infty \tag{17}$$

Asymptotic normality

The distribution of estimators is said to be asymptotically normal if, as the sample size increases, the distribution of the estimators approaches a normal distribution. To show that our estimators are asymptotically normal we need the following additional assumptions.

D_7. σ_{*t}^2 is twice continuously differentiable on subspace Θ_1

D_8. The following moment conditions hold: $E(e_0^4) < \infty$, $E\left[\frac{|\nabla\sigma_{*0}^2(\theta_0)|^2}{\sigma_0^4}\right] < \infty$, $E\|\nabla l_n\|_{\Theta_1} < \infty$ and $E\|\nabla^2 l_n\|_{\Theta_1} < \infty$.

If the conditions $D_1 - D_8$ hold, the following theorem can be stated for the asymptotic normality of the quasi-Maximum likelihood estimator.
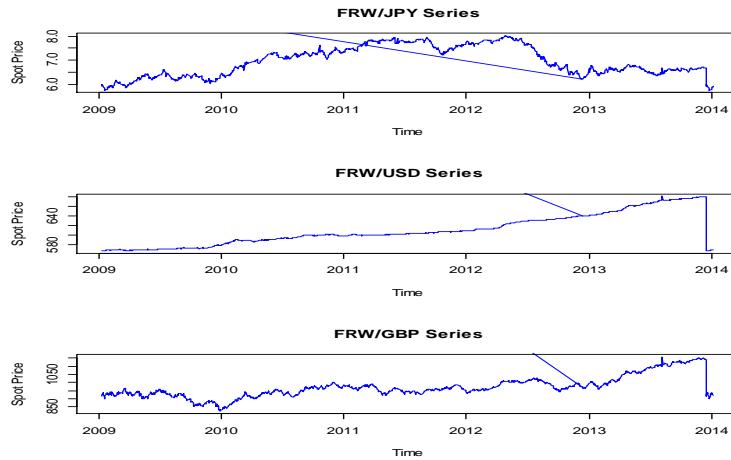
Theorem

Under the conditions $D_1 - D_8$, the Quasi-Maximum Likelihood Estimator $\hat{\theta}_n$ is strongly consistent and asymptotically normal, that is

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{d} N(0, \; V_0) \text{ as } n \to \infty. \tag{18}$$

Where V_0 is the asymptotic variance of the estimator $\hat{\theta}_n$. Under asymptotic normality, the estimator $\hat{\theta}_n$ not only converges to the unknown parameter, but it converges fast enough, at a rate $1/\sqrt{n}$. For more details see (Francq and Zakoïan, 2004) and (Posedel, 2005).
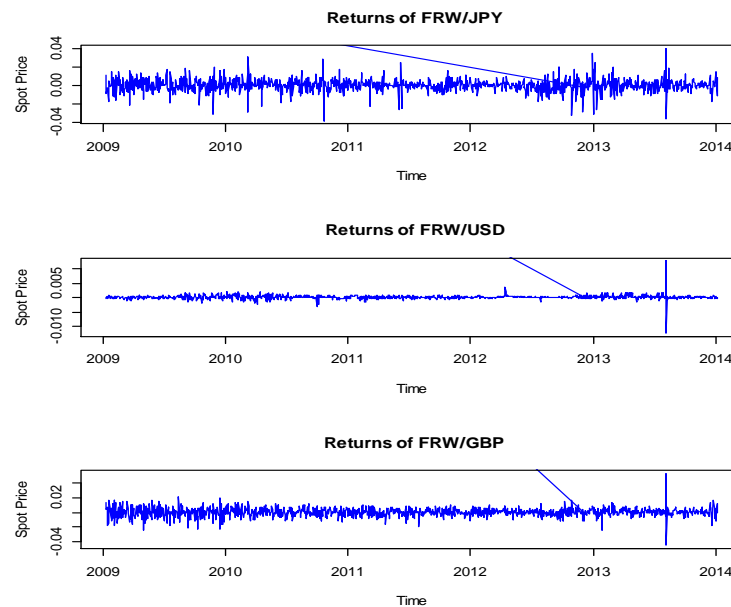
EMPIRICAL RESULTS

This chapter presents empirical results on the estimation of extreme risk in exchange rates using volatility and Extreme value theory. The plots in Figures 1 below show the daily fluctuations of Exchange rate series.

Figure 1: Trends in the Exchange rates

In order to estimate the volatility in the exchange rates data, the Log-returns was applied. The log-returns in plots in Figure 2 show that the data appear to be stationary in mean after logarithm transformation. These plots also reveal dependence structure where period of high returns tend to be followed by the high returns as well as the period of low returns tend to be followed by low returns. This is evidence of short-range dependence (volatility clustering in data), which must cast doubt on the assumption of independent and identically distributed $i.i.d$ of data. The clustering of exchange rate returns data indicates presence of stochastic volatility in exchange rate series. Plots in Figure 2 allow identifying the most extreme losses and their approximate time occurrence.

Figure 2: Daily Exchange rate returns

Descriptive statistics for the exchange rate returns are presented in Table 1. The mean of the exchange rate returns is negligible for each currency but standard deviation is significant for all currencies. The distributions of returns in FRW/JPY and in FRW/GBP exhibit negative skewness (means frequent small gains and few extreme losses). This indicates that they have what

29

statisticians call a long left tail, which for investors can mean a greater change of extremely negative outcomes. However the returns series of FRW/USD exhibit positive skewness coefficients. This indicates that the distribution of the returns in this currency is slightly right skewed. This implies that depreciations in exchange rates occur slightly more often than appreciation. In this case investors can have frequent small negative outcomes and few extreme gains.

The excess kurtosis coefficients for log-returns of all currencies are positive to indicate that the underlying distributions of exchange rate returns have tails which are significantly heavier than that of the normal distribution. Jarque Bera test for normality rejects null hypothesis for all currencies since $p - value$ is too small compared to significance level. This indicates that none of these currencies is normally distributed. Augmented Dickey Fuller (ADF) test used for stationarity testing revealed that exchange rate returns for all currencies are stationary in mean. The ADF statistics used in the test is a negative number for all the currencies. More negative indicates the stronger the rejection of the null hypothesis.

Table 1: summary statistics of Log-Returns data

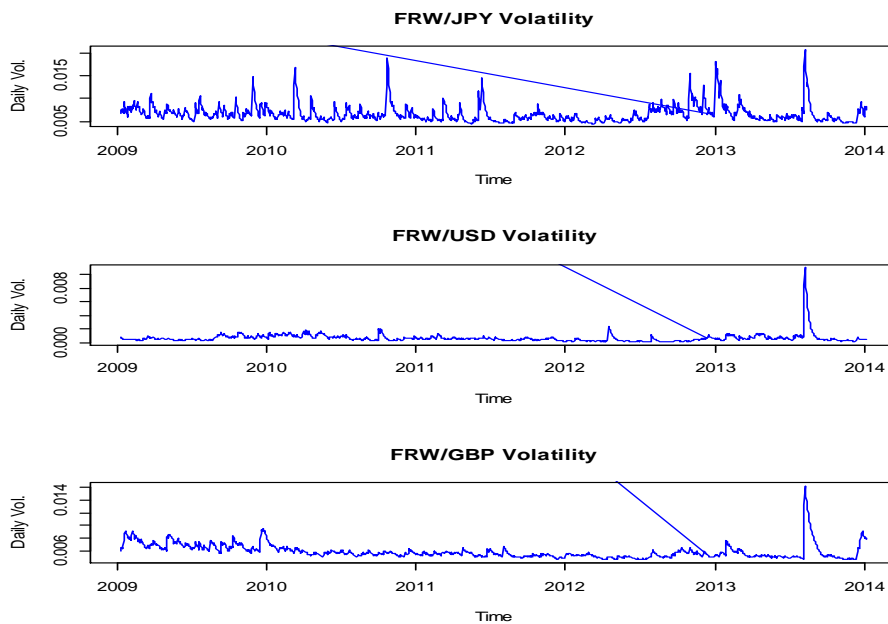| Statistics	FRW/JPY	FRW/USD	FRW/GBP
Mean	9.7e-05	0.00015	0.0001729
Minimum	-0.0386	-0.01119	-0.04536
Maximum	0.03989	0.008339	0.053518
Standard dev.	0.00676	0.000746	0.005989
Skewness	-0.3285	0.050277	-0.07602
Excess.Kurtosis	5.66189	129.5456	8.307359
JB.Test(p-value)	<2.2e-16	< 2.2e-16	< 2.2e-16
ADF. Statistics(10)	-10.753	-8.8158	-10.4837
ADF.Test(p-value)	0.01	0.01	0.01
Significance level	0.05	0.05	0.05
No. of Observations	2758	2758	2758

The Autocorrelation function (ACF) and Partial Autocorrelation function (PACF) were applied to obtain the lags in the GARCH (p, q) model. These functions help to determine which past series values are most useful in predicting future values. The length of past conditional variance (p) was determined by ACF where the lag at which the ACF cuts off is the indicated the number of GARCH term (p). The PACF determine the length of past squared innovations (q) where the lag at which PACF cuts off is the indicated number of ARCH term (q). The Akaike information (AIC) criterion and Bayesian information criterion (BIC) tests were applied to select the best model for each currency. As suggested by Akaike (1973) and Schwarz (1978), the best model for financial data is the one that minimize the AIC and BIC respectively. The results presented in Table 2 show that GARCH (1, 1) is the best model for all currencies.

Table 2: summary of GARCH model selection

Frw/	$\widehat{\omega}$	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\alpha} + \widehat{\beta}$	observations
JPY	4.93e-06	0.1569	0.7463	0.9032	Low persistent
USD	7.17e-09	0.4383	0.7403	1.1786	High persistent
Pound	2.2e-06	0.0529	0.885	0.9379	Low persistent

The plots in Figure3 reveal volatility clustering characteristics. Statistically, volatility clustering implies a strong autocorrelation. Volatility clustering describes the tendency of large changes in exchange rate series to follow large changes and tendency of low changes to be followed by low change. In other word the current level of volatility tends to be positively correlated with the preceding periods. The sum of the ARCH and GARCH coefficients for FRW/USD exceed unity. This indicates that the conditional variance is highly persistent to the shocks in the volatility of the Frw/USD currency; therefore, the memory of the shocks for this currency is remembered in the exchange rates markets. The sum of ARCH term and GARCH term coefficients for Frw/JPY and for Frw/Pound are below the unity. This indicates that the variance is relatively less persistent of the shocks in the volatility.

Figure 3: Daily exchange rate volatility

In all GARCH models, the estimated parameters are significant at 5% level. Lagrange Multiplier (LM) test for ARCH effects rejects the null hypothesis of no ARCH effects at 12 degrees of freedom for Frw/JPY. This indicates the presence of ARCH effects in residuals of exchange rate returns. The ARCH effects decrease with the increase of the number of lags. The ARCH effects in the exchange rate returns for Frw/JPY dies out after (64) lags. There is no arch effect in the Frw/USD and in Frw/GBP at 12 degrees of freedom. LM test shows that there is no Arch effect in Frw/USD at all lags however the Arch effect in Frw/GBP dies out after (9) lags.

CONCLUSION AND RECOMMENDATION

In this work, the GARCH models were applied to estimate volatility in the Rwandese foreign exchange market data for the period June 2009 to June 2014. Quasi-Maximum Likelihood estimation procedure was used to estimate parameters in GARCH and the estimators were found to be consistency and asymptotically normal. The exploratory analysis showed that exchange rate data exhibit leptokurtosis. Lagrange Multiplier test for ARCH effects revealed presence of ARCH effects in both returns series and residuals. Augmented Dickey Fuller test for unit roots showed that the return series for all currencies are stationary in mean. Results of this work will contribute a lot to understanding of how changes in exchange rate affect the prices of goods and services in international trade. Nelson (1991) introduced Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) model and listed three shortcomings with the GARCH models; the lack of symmetry in response to shocks, the GARCH models impose parameter restrictions to ensure the conditional variance is positive and the difficulty in measuring the persistence using GARCH model. From these drawbacks, it is recommended that the future research should focus on asymmetric models to see whether these shortcomings can have significance impact to volatility estimation.

References

Akaike, H. (1973). Information Theory and Extension of the Maximum Likelihood Principle. . *In: B.N. Petrov and F.Cs'aki Eds. Akad emia kiad'o, Budapest* , 267-281.

Bera, A., & Jarque, C. (1982). Model Specification Tests: A simultaneous Approach. *Journal of Econometrics Vol.20* , 57-82.

Blum, P., & Dacorogna, M. (2002). Extreme moves in daily Foreign Exchange rates and risk limit setting. . *Department of Mathematics General Guisan Quasi Vol26* , , 8092Z'urich.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity . *Journal of Econometris Vol.31* , 307-327.

Engle, F. (19982). Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation. *Econometrica Vol.50* , 987-1008.

Francq, C., & Zakoian, J. (2004). Maximam Likelihood Estimationof Pure GARCH and ARMA-GARCH Processes. *Bernoulli Vol.10* , 605-637.

Ghysels, E., Harvey, A., & Renault, E. (1996). Stochastic Volatility. In Handbook of Statistics. *Statistical Methods in Finance, Maddala, G.S. Ed. North- Holland, Amsterdam Vol.14* , 119-91.

Glosten, L., Jagannathan, R., & Runkle, D. (1993). On the relation between the expected value and the Volatility of Nominal excess return on stocks. *Journal of Finance Vol.48* , 1779-1801.

Hull, J., & White, A. (1998). Value at Risk When Changes in Market Variables are not Normally Distributed. *Journal of Risk Vol.1* , 47-61.

Longin, F. (1996). The Asymptotic Distribution of Extreme Stock Market Returns . *Journal of Business Vol.67* , 383-408.

Maan, I., Mwita, N., & Odhiambo, R. (2010). Modelling the Volatility of Exchange Rates in the Kenyan Markets. *Journal of Business Management Vol.4* , 1401-1408.

Manganelli, S., & Engle, R. (2001). Value at Risk Models in Finance,European Central Bank Working Paper Series, Frankfurt. Modelling. *Mathematical Finance Vol.14* , 75-102.

Neftci, S. (2000). Value at Risk Calculations, Extreme Events and Tail Estimation. *Journal of Derivatives Vol.7* , 23-38.

Nelson, D. (1991). Conditional Heteroscedasticity in asset returns: A new Approach. *Econometrica Vol.59* , 347-370.

Posedel, P. (2005). Properties and Estimation of GARCH(1,1) Model . *Metodolo S Ki Zvezki Vol.2* , 243-257.

Sandmann, G., & Koopman, S. (1998). Estimation of Stochastics Volatility Models Via Monte Carlo Maximum Likelihood . *Journal of Econometrics Vol.87* , 271-301.

Schwartz, G. (1978). Estimating the Dimension of a Model . *Annals of Statistics Vol6.* , 461-464.

Smith, W., Clifford, C., & Stykes, W. (1990). Managing Financial Risk . *New York, Harper and Row.*