International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

Modeling Volatility Using Garch Models: Application to Food Inflation Volatility in Rwanda

Placide Aime Kwizera

Placide Aime Kwizera is an economist and Invited lecturer of economics at Kigali Independent University

Abstract: ***The aim of this paper is to introduce GARCH-modelling and its application on Rwanda food inflation data spanning from January, 2004 to December, 2018(180-observations). On the basis of estimation results of various GARCH models and diagnostic check has shown that the AR-GARCH with Gaussian distributed innovations is most appropriate specification for modeling food inflation volatility in Rwanda. The study finds no evidence of asymmetry in the response of food inflation volatility to negative and positive shocks. Simulation on estimated AR-GARCH with Gaussian distributed innovations and a 6-years (72months) forecast from January 2019 to December 2024 were also made. Hence, the AR-GARCH with Gaussian distribution of innovations could be a widely useful tool for modelling the food inflation volatility in Rwanda.***

Keywords: Inflation; volatility; GARCH; Leverage effect

JEL Classification: C22, C51, C52, E31

1. Introduction

Time series analysis is an ordered sequence of values of a variable at equally spaced time intervals. It is used to understand the determining factors and structure behind the observed data, choose a model to forecast, thereby leading to better decision making (Chatfield, C. 2000). Inflation as described by Webster, D. (2000) is the persistent increase in the level of consumer prices or persistent decline in the purchasing power of money. Currently, inflation is the one of the major economic challenges in the most developing countries and it becomes focus of economic policy worldwide as described by (David, F.H (2001). The inflation dynamics can be studied using a stochastic modelling approach that captures the time dependent structure embedded in the time series inflation data.

The autoregressive conditional heteroscedasticity (ARCH) models, with its extension to generalized autoregressive conditional heteroscedasticity (GARCH) models as introduced by Engle, R. (1982) and Bollerslev, T. (1986) respectively accommodate the dynamics of conditional heteroscedasticity have been found to be useful for several time series modelling, the applications in economics and finance have been particularly successful.

This study introduce GARCH-modelling and its application on Rwanda food inflation data to choose the most suitable model that explains behavior of food inflation in Rwanda.

The rest of this paper is structured into three sections after the introduction; GARCH modeling is presented in section two while the third section presents application of GARCH modeling to Rwanda food inflation data.

2. Garch Modelling

2.1. Overview

The stylized characteristics of financial time series data include: almost zero correlation, absolute/squared data exhibit high correlation and excess kurtosis/ heavy tailed distribution. Consider the general form of conditional variance model

$$y_t = \mu_t + e_t \quad \text{.................................(1)}$$

$$e_t = \sigma_t \varepsilon_t \quad \text{...................................(2)}$$

Firstly we see that value of dependent variable y_t consists of mean μ_t and innovation e_t. In practice μ_t can be chosen as conditional mean of y_t such that $\mu_t = \mathrm{E}(y_t / \Omega_{t-1}) = g(\Omega_{t-1})$ where Ω_{t-1} is arbitrary historical information affecting value of y_t. In other words we model every μ_t by suitable linear regress model or using AR process. Often is sufficient use just fixed value μ_t = μ or μ_t =0. Innovation ε_t consists of variance (volatility) root σ_t where $\sigma_t^2 = h_t = h(\Omega_{t-1}) = \mathrm{var}(e_t / \Omega_{t-1}) = \mathrm{var}(y_t / \Omega_{t-1})$ and *i.i.d.* random variable from normal or t-distribution $\varepsilon_t \sim N(0,1)$ or $\varepsilon_t \sim t(v)$.

Originally, variance (volatility) is developed based on Standard Normal (Gaussian) distribution .In other words $\eta_t = \varepsilon_t / \sqrt{h_t} \sim N(0,1)$. Hence, the conditional density of ε_t can be constructed as follows.

$$\Omega_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_1\}$$

$$\theta(\mu, \omega, \alpha, \beta)'$$

$$f(\varepsilon_t / \theta, \Omega_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} e^{\frac{\varepsilon_t^2}{2h}} \quad \text{....(3)}$$

Then, the log-likelihood function corresponding to equation (3) is:

$$L(\theta / \varepsilon) = \sum_{t=2}^{T} \ln f(\varepsilon_t / \theta, \Omega_{t-1}) \quad \text{.................(4)}$$

where

$\varepsilon = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_T)'$

and MLE estimator $\hat{\theta}$ is obtained by maximizing Equation (4). Additionally, standard deviation of $\hat{\theta}$ is acquired by taking square root of diagonal terms of the inversed fisher information (Bollerslev, 1986)

However, significant evidence suggests that the financial time series is rarely Gaussian but typically leptokurtic and heavy-tailed (see, for example, Bollerslev, 1987). Therefore, if the true distribution is not Gaussian, MLE standard deviation of $\hat{\theta}$ estimated in the above procedure will be inconsistent. To solve this problem, the Quasi-Maximum Likelihood Estimation (QMLE) based on Gaussian is further derived. The algorithm of QMLE to estimate $\hat{\theta}$ is the same as described that by Equations (3) and (4). The only difference is the way to estimate a robust standard deviation of $\hat{\theta}$ (Bollerslev & Wooldridge, 1992). It is argued that the QMLE standard deviation is asymptotically consistent, even if the true distribution of η_t is not Gaussian.

Alternatively, although QMLE can lead to consistent estimates, it is argued that QMLE of GARCH model is not efficient. Among the existing literature, Student's t-distribution is widely used alternative in finance research (Chkili et al., 2012; Fan et al., 2008; Mabrouk & Saadi, 2012; 2011). This distribution can capture leptokurtic and heavy-tail behaviours. When it applied to the variance (volatility) , the corresponding density functions of ε_t is described below.

$$f(\varepsilon_t \backslash \theta, \Omega_{t-1}) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{\pi(v-2)h_t}} \left[1 + \frac{\varepsilon_t^2}{(v-2)h_t}\right]^{-\frac{v+1}{2}}$$

.................(5)

The MLE estimator $\hat{\theta}$ can be obtained in the same way as that described on normal distribution.

From fragmented notation above we can write general conditional variance model as is known in econometrics literature

$$y_t = \mu_t + \sigma_t \varepsilon_t = \mu_t + \sqrt{h_t} \varepsilon_t = g(\Omega_{t-1}) + \sqrt{h(\Omega_{t-1})} \varepsilon_t$$

.................(6)

Observe that innovations e_t are not correlated but are dependent through σ_t term (later we will briefly see Ω_{t-1} contains lagged σ_t^2). If g is non-linear function then model is non-linear in mean, conversely if h is non-linear then model is non-linear in variance which in turn means that h_t is changing non-linearly with every t through a function of Ω_{t-1}. Since now we should know what autoregressive conditional heteroskedasticity means.

2.2. ARCH Model

The ARCH model which was introduced by Rober Engle in 1982. If we take (6) and specify condition (based on historical information Ω_{t-1}) for σ_t we get ARCH(m) model

$y_t = \mu_t + e_t$ where $e_t = \sigma_t \varepsilon_t$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i e_{t-1}^2 \quad \text{..............................(7)}$$

where ε_t are i.i.d. random variables with normal or t-distribution, zero mean and unit variance. As mentioned earlier in practice we can drop μ_t term thus get

$y_t = e_t$, where $e_t = \sigma_t \varepsilon_t$,

$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i e_{t-1}^2$ or regress mean with exogenous explanatory variables as

$y_t = \gamma_0 + \sum_{i=1}^{k} \gamma_i x_{t,i} + e_t$, where $e_t = \sigma_t \varepsilon_t$,

$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i e_{t-1}^2$ or use any other suitable model.

Recall that significant fluctuation in past innovations will notably affect current volatility (variance). Regarding positivity and stationarity of variance σ_t^2 , coefficients in (7) condition have to satisfy following constraints

$$\alpha_0 \succ 0, \alpha_1 \geq 0, ..., \alpha_m \geq 0, + \sum_{i=1}^{m} \alpha_i \prec 1$$

Properties of ARCH (m) Model

The Mean: From equation 7the conditional expectation and variance of is: x_t $\mathrm{E}(y_t) = 0$ since the expectation of $\varepsilon_t = 0$

The Second Moment or Variance:

$\mathrm{E}(y_t^2) = \mathrm{E}(\sigma_t^2 \varepsilon_t^2) = \mathrm{E}(\sigma_t^2)$,since $\sigma^2 = 1$ following a standard normal distribution of ε_t

$$\mathrm{E}(\sigma_t^2) = \alpha_0 + \alpha_i \sum_{i=1}^{m} \mathrm{E}(y_{t-1}^2), \quad \text{given } \mathrm{E}(\sigma_t^2) = \mathrm{E}(y_{t-1}^2)$$

under stationarity assumption,

$$\mathrm{E}(\sigma_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{m} \alpha_i}$$

For ARCH(1), the variance is given by:

$$\mathrm{E}(\sigma_t^2) = \frac{\alpha_0}{1-\alpha_1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8)$$

The Kurtosis:

First, the forth moment of the time series is obtained,

$$\mathrm{E}(y_t^4) = \mathrm{E}\left\{(\sigma_t^2)^2 \varepsilon_t^4\right\} = \mathrm{E}\left\{(\sigma_t^2)^2\right\}\mathrm{E}(\varepsilon_t^4) = 3\mathrm{E}\left\{(\sigma_t^2)^2\right\}$$

$$\mathrm{E}\left\{(\sigma_t^2)^2\right\} = \mathrm{E}\left\{(\alpha_0 + \sum_{i=1}^{m}\alpha_i y_{t-1}^2)^2\right\} = \alpha_0^2 + 2\alpha_0\sum_{i=1}^{m}\alpha_i \mathrm{E}(y_{t-}^2$$

Substituting equation, we have:

$$\mathrm{E}(y_t^4) = 3\left\{\alpha_0^2 + 2\alpha_0\sum_{i=1}^{m}\alpha_i \mathrm{E}(y_{t-1}^2) + \sum_{i=1}^{m}\alpha_i^2 \mathrm{E}(y_{t-1}^4)\right\}$$

Under stationarity,

$$\mathrm{E}(y_t^2) = \mathrm{E}(\sigma_t^2) = \frac{\alpha_0}{1-\sum_{i=1}^{m}\alpha_i} \text{ and } \mathrm{E}(y_t^4) = \mathrm{E}(y_{t-1}^4)$$

$$\mathrm{E}(y_t^4) = 3\frac{\alpha_0^2(1+\sum_{i=1}^{m}\alpha_i)}{(1-\sum_{i=1}^{m}\alpha_i)(1-3\sum_{i=1}^{m}\alpha_i^2)}$$

……………………………………………………….……(9)

The Kurtosis is given by: $K(y) = \dfrac{\mathrm{E}(y_t^4)}{\left\{\mathrm{E}(y_t^2)\right\}^2}$

Substituting equations (8) and (9), we get:

$$K(y) = 3\frac{(1+\sum_{i=1}^{m}\alpha_i)(1-\sum_{i=1}^{m}\alpha_i)}{1-3\sum_{i=1}^{m}\alpha_i^2}$$,therefore, the kurtosis

is $K = 3\dfrac{1-\sum_{i=1}^{m}\alpha_i^2)}{1-3\sum_{i=1}^{m}\alpha_i^2}$

When $i=1$, we get ARCH(1), then the Kurtosis of ARCH(1) is;

$$K = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2}$$

…………………………………………………………………
………………….(10)

Which is strictly greater than 3 unless $\alpha_1 = 0$ .The kurtosis for a normally distributed random variable Z is 3. Thus, the kurtosis of y_t is greater than the kurtosis of a normal distribution, and the distribution of y_t has a heavier tail than the normal distribution, when $\alpha_1 \succ 1$.

Fitting Procedure for ARCH model

There are two steps in model fitting:

Step1: Plotting the inflation series and analyzing the autocorrelation function (ACF) and the partial autocorrelation function (PACF). We check for correlation in the inflation series by performing the autocorrelation function to compute and display the sample ACF of the returns and by plotting the partial correlation functions. The ACF definition (Auto correlation function)

A time series $\{X_t\}$ has mean function $\mu_t = \mathrm{E}[X_t]$

The auto correlation function (ACF) is

$$\rho x(h) = \frac{\gamma x(h)}{\gamma x(0)} = corr(X_{t+h,} X)$$

PACF Definition (Partial Autocorrelation Function) of the k-th order is defined as:

$$P_{kk} = \varphi_{kk} = corr(X_t - P(X_t \backslash X_{t+1,\ldots}X_{t+k-1})X_{t+k} - P(X_{t+k} \backslash X_{t+1,\ldots,}X_{t+k-1}))$$

Step2: Performing preliminary tests, such as ARCH effect test or the Q-test.

We can quantify the preceding qualitative checks for correlation using formal hypothesis checks, like Ljung-Box-Pierce Q-test and Engle's ARCH test. By performing a Ljung-Box-Pierce Q-test, we can verify, at least approximately, the presence of any significant correlation in the inflation when tested for up to 20 lags of the ACF at the 0.05 level of significance.

Weakness of ARCH model

Despite ARCH model able to capture the characteristics of financial time series data, it has some weaknesses that may make GARCH model better. These weaknesses include; ARCH treats positive and negative inflation in the same way (by past square inflation), it is very restrictive in parameters, it does not provide any new insight for understanding financial time series, it often over-predicts the volatility, because it respond slowly to large shocks and volatility from it persists for relatively short amount of times unless m is large.

2.3. GARCH Model

GARCH model was introduced by Robert Engle's(1982) and Tim Bollerslev in 1986. Both GARCH and ARCH models allow for leptokurtic distribution of innovations e_t and volatility clustering (conditional heteroskedasticity) in time series but neither of them adjusts for leverage effect. So what is the advantage of GARCH over ARCH? ARCH model often requires high order m thus many parameters have to be estimated which in turn brings need for higher computing power. Moreover the bigger order m is, the higher probability of breaking aforementioned constraints there is. GARCH is "upgraded" ARCH in that way it allows

ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

current volatility to be dependent on its lagged values directly. GARCH(m , n) is defined as

$y_t = \mu_t + e_t$, where $e_t = \sigma_t \varepsilon_t$,

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i e_{t-i}^2 + \sum_{i=1}^{m} \beta_i \sigma_{t-j}^2 \text{ ....(11)}$$

Where ε_t are i.i.d. random variables with normal or t-distribution, zero mean and unit variance. Parameters constraints are very similar as for ARCH model,

$$\alpha_0 \succ 0, \alpha_i \geq 0, \beta_j \geq 0, \sum_{i=1}^{m} \alpha_i + \sum_{j=1}^{n} \beta_j \prec 1$$

In practice even GARCH (1,1) with three parameters can describe complex volatility structures and it's sufficient for most applications. We can forecast future volatility $\hat{\sigma}_{t+\tau}^2$ of GARCH(1,1) model using

$$\hat{\sigma}_{t+\tau}^2 = \sigma^2 + (\alpha_1 + \beta_1)^\tau (\sigma_t^2 + \sigma^2)$$

Where, $\sigma^2 = \dfrac{\alpha_0}{1 - \alpha_1 - \alpha_2}$

is unconditional variance of innovations e_t. Observe that for $\alpha_1 + \beta_1 \prec 1$ as $\tau \to \infty$ we get $\hat{\sigma}_{t+\tau}^2 \to \sigma^2$. So prediction of volatility goes with time asymptotically to the unconditional variance.

Properties of GARCH (m,n)

The mean:
From equation (3), the conditional expectation and variance of x_t is: $\mathrm{E}(y_t) = 0$ since the expectation of $\varepsilon_t = 0$

The Second Moment or Variance:

$$\mathrm{E}(y_t^2) = \mathrm{E}(\sigma_t^2 \varepsilon_t^2) = \mathrm{E}(\sigma_t^2)$$

$$\mathrm{E}(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{m} \alpha_i \mathrm{E}(y_{t-i}^2) + \sum_{j=1}^{m} \beta_j \mathrm{E}(\sigma_{t-j}^2)$$

Given $\mathrm{E}(\sigma_t^2) = \mathrm{E}(y_{t-1}^2) = \mathrm{E}(\sigma_{t-j}^2)$ under stationarity assumption,

$$\mathrm{E}(\sigma_t^2) = \frac{\alpha_0}{1 - (\sum_{i=1}^{m} \alpha_i + \sum_{i=1}^{n} \beta_i)}$$

For GARCH (1, 1)

$$\mathrm{E}(\sigma_t^2) = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)} \text{ ....(12)}$$

The Kurtosis:
First the forth moment of the time series is obtained;

$$\mathrm{E}(y_t^4) = \mathrm{E}\left\{(\sigma_t^2)^2 \varepsilon_t^4\right\} = \mathrm{E}\left\{(\sigma_t^2)^2\right\} \mathrm{E}(\varepsilon_t^4) = 3\mathrm{E}\left\{(\sigma_t^2)^2\right\}$$

but $\mathrm{E}\left\{(\sigma_t^2)^2\right\} = \mathrm{E}\left\{(\alpha_0 + \sum_{i=1}^{m} \alpha_i y_{t-i}^2 + \sum_{j=1}^{n} \beta_j \sigma_{t-j}^2)^2\right\}$

$$\alpha_0^2 + 2\alpha_0 \sum_{i=1}^{m} \alpha_i \mathrm{E}(y_{t-i}^2) + 2\alpha_0 \sum_{j=1}^{n} \beta_j \mathrm{E}(\sigma_{t-j}^2) + \sum_{i=1}^{m} \alpha_i^2 \mathrm{E}(y_{t-i}^4)$$

$$+ \sum_{j=1}^{n} \beta_j^2 \mathrm{E}\left[(\sigma_{t-j}^2)^2\right] + 2\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j \mathrm{E}(y_{t-i}^2 \sigma_{t-j}^2)$$

=When $i = j = 1$, we get GARCH (1, 1)

$$\mathrm{E}\left\{(\sigma_t^2)^2\right\} = \alpha_0^2 + \alpha_1^2 \mathrm{E}(y_{t-1}^4) + \beta_1^2 \mathrm{E}\left\{(\sigma_{t-1}^2)^2 + 2\alpha_1 \beta_1 \mathrm{E}(y_{t-1}^2 \sigma_{t-1}^2) + 2\alpha_0 \alpha_1 \mathrm{E}(y_{t-1}^2) + 2\alpha_0 \beta_1\right\} \mathrm{E}(\sigma_{t-1}^2)$$

Assuming the process is stationary, $\mathrm{E}\left\{(\sigma_t^2)^2\right\} = \mathrm{E}\left\{(\sigma_{t-1}^2)^2\right\}$ Hence

$$\mathrm{E}\left\{(\sigma_t^2)^2\right\} = \frac{\alpha_0^2 + 2\alpha_0(\alpha_1 + \beta_1)\mathrm{E}(\sigma_{t-1}^2)}{1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_t^2}$$

$$= \frac{\alpha_0^2 + 2\alpha_0^2(\alpha_1 + \beta_1)}{(1 - \alpha_1 - \beta_1)(1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_t^2)}$$

$$\mathrm{E}(y_t^4) = 3\mathrm{E}\left\{(\sigma_t^2)^2\right\}$$

$$K = 3\frac{\alpha_0^2 + 2\alpha_0^2(\alpha_1 + \beta_1)}{(1 - \alpha_1 - \beta_1)(1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2)} \text{ ....(13)}$$

The Kurtosis is given by; $K = \dfrac{\mathrm{E}(y_t^4)}{\left\{\mathrm{E}(y_t^2)\right\}^2}$

Substituting equation (12) and equation (13), we get;

$$K = 3\frac{1 - (\alpha_1 + \beta_1)}{1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2)} \text{ ....…….(14)}$$

Which is strictly greater than 3 unless $\alpha_1 = 0$

2.4. GJR-GARCH Model

There are some aspects of the model which can be improved so that it can better capture the characteristics and dynamics of a particular time series in leverage effects, volatility clustering and leptokurtosis are commonly observed in financial time series. The model which adjusts even for asymmetric responses of volatility to innovation fluctuations. GJR-GARCH was developed by Glosten, Jagannathan, Runkle in 1993. Sometimes referred as T-GARCH or TARCH if just ARCH with GJR modification is used. GJR-GARCH(p,q,r) is defined as follows

$y_t = \mu_t + e_t$, where $e_t = \sigma_t \varepsilon_t$

$$\sigma^2{}_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i e_{t-i}^2 + \sum_{i=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{i=k}^{r} \gamma_k e_{t-k}^2 I_{t-k}, I_t \begin{cases} 1 \text{ if } e_t < 0 \\ 0 \text{ otherwise} \end{cases} \text{ ....(15)}$$

where γ_k are leverage coefficients and I_t is indicator function. Observe that for $\gamma_k \succ 0$ negative innovations e_t give additional value to volatility σ_t^2 thus we achieve adjustment for asymmetric impact on volatility as discussed at the beginning of the article. For $\gamma_k = 0$ we get GARCH (m = p, n = q) model and for $\gamma_k \prec 0$ we get exotic result where upward swings in return or price have stronger impact on volatility than the downward moves. Need to mention that in most implementations of GJR-GARCH we will find GJR-GARCH (p,q) where leverage order 0 is automatically considered equal to order p. Parameters constraints are again very similar as for GARCH, we have

$$\alpha_0 \succ 0, \alpha_i \geq 0, \beta_j \geq 0, \alpha_i + \gamma_k \geq 0, \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j \frac{1}{2} \sum_{k=1}^{r} \gamma_k \prec 1$$

Prediction for GJR-GARCH can be estimated as

$$\hat{\sigma}_{t+\tau}^2 = \alpha_0 + \frac{\alpha_1 + \gamma_1}{2} + \beta_1 \sigma_t^2 (\tau - 1) \ldots\ldots\ldots\ldots\ldots(16)$$

2.5. Model selection criteria

Selection criteria assess whether a fitted model offers an optimal balance between the goodness-of-fit and parsimony. The most common model selection criteria such as the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the Schwarz information criterion (SIC), Hannan Quinn Criterion (HQ) and Loglikelihood (LL)were used as bases for selection criteria.

- $AIC = 2\log(\max likelihood) + 2k$ where, $2k = p + q + 1$ if the model contains an intercept or a constant term and $k = p + q$
- $BIC = -2\log(L) + 2(m)$
- $HQ = -2\log(L) + 2m\log(\log n)$

Where m and n are number of observations (sample size) and parameter in the model respectively and log l is the $\log$ likelihood. The desirable model is one that minimizes the AIC, the BIC, the HQ, SIC and LL.

2.6. Forecast of Conditional Variance in GARCH model

The formula used to calculate the multi-step ahead forecasts of the conditional variance for the GARCH(1,1)model is obtained as illustrated. For such a model, the variance equation is

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \ldots\ldots\ldots(17)$$

Denote the forecast origin by n and the forecast horizon by h let F_n be the information set available at time n. For $h = 1$, the 1-step ahead forecast of the conditional variance is simply

$$\mathrm{E}(\sigma_t^2 \backslash F_n) = \mathrm{E}(\alpha_0 + \alpha_1 y_n^2 + \beta_1 \sigma_n^2 \backslash F_n) = \alpha_0 + \alpha_1 y_n^2 + \beta_1 \sigma_n^2 \ldots\ldots\ldots\ldots(18)$$

For $h = 2$ by using the assumption that Z_t' s are i.i.d. $N(0,1)$, we have

$$\mathrm{E}(\sigma_{n+2}^2 \backslash F_n) = \mathrm{E}(\alpha_0 + \alpha_1 y_{n+1}^2 + \beta_1 \sigma_{n+1}^2 \backslash F_n) = \alpha_0 [1 + (\alpha_1 + \beta_1)] + \alpha_1 (\alpha_1 + \beta_1) y_n^2 + \beta_1 (\alpha_1 + \beta_1) \sigma_n^2$$
$$= \alpha_0 + (\alpha_1 + \beta_1) * \mathrm{E}(\sigma_{n+1}^2 \backslash F_t) \ldots\ldots\ldots\ldots\ldots\ldots\ldots(19)$$

By the same argument, it is easily seen that for $j = h$, the j-step ahead forecast of the conditional variance of the GARCH(1,1) model is

$$\mathrm{E}(\sigma_{n+j}^2 \backslash F_n) = \alpha_0 \sum_{k=0}^{j-1} (\alpha_1 + \beta_1)^k + \alpha_1 (\alpha_1 + \beta_1)^{j-1} y_n^2 + \beta_1 (\alpha_1 + \beta_1)^{j-1} \sigma_n^2 = \alpha_0 [1 + (\alpha_1 + \beta_1)] + \alpha_1 (\alpha_1 + \beta_1) y_n^2 + \beta_1 (\alpha_1 + \beta_1) \sigma_n^2$$
$$= \alpha_0 + (\alpha_1 + \beta_1) * \mathrm{E}(\sigma_{n+j-1}^2 \backslash F_t) \ldots\ldots\ldots\ldots(20)$$

Therefore, the forecasts of the conditional variances of an GARCH (1,1) model can be computed recursively.

3. Application Grach Modeling to Application to Food Inflation Volatility in Rwanda

3.1 Introduction

Modelling food inflation volatility is crucial for the policy makers, it provides path of policy formulation particularly for central bank to achieve the price stability. The price instability can generally jeopardize the entire macroeconomic stability (Bonato, 1998).

It is vital for central bank to understand the future path of inflation to anchor expectations and ensure policy credibility; the key aspects of an effective monetary policy transmission mechanism (King, 2005).

We apply GARCH models on food inflation volatility to choose the most suitable model that explains behavior of food inflation in Rwanda using dataset spanning from January, 2004 to December, 2018.

3.2 Brief empirical review

Ngailo and Massawe (2014) used monthly inflation data observations from Tanzania and considered the GARCH approach in modelling inflation rates for eleven years from 2000 to 2011. After performing all the diagnostic checks on Jarque bera test on kurtosis and the stationarity using Augmented ducker fuller test. They found out that the inflation returns volatility works better with the class of GARCH (1, 1).

Awogbemi and Oluwaseyi (2011) results showed that ARCH and GARCH models are better models because they give lower values of AIC and BIC as compared to the conventional Box and Jenkins ARMA models for inflation in Nigeria. The researchers also observed that since volatility seems to persist in all the commodity items, people who expect a rise in the rate of inflation (the 'bullish crowd') will be highly favored in the market of the said commodity items.

Jiang (2011) believed that it was worthy to investigate the inflation and inflation uncertainty relationship in China as it is commonly believed that one possible channel that inflation imposes significant economic costs is through its effect on inflation uncertainty. He addressed the relationship of inflation and its uncertainty in China's urban and rural areas separately given the huge urban-rural gaps In conclusion he said that TGARCH (1,1) was the best in studying the inflation rate volatility in China.

An important characteristic of Rwanda's food inflation data is volatility which has been measured in term of food CPI in log difference as shown in the following:

4. Results and Discussions

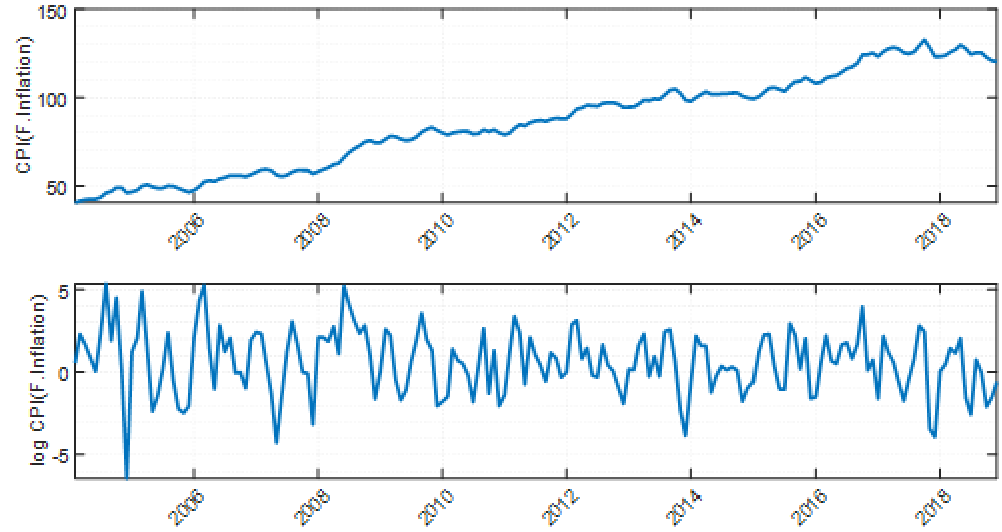

Figure 1: Volatility in food inflation series
Source: Author's computation

The mean reverting (food inflation tend to remain around a certain value) property can also be seen clearly where the food inflation revolve around zero. Let's examine character of food inflation mean, ACF, PACF and Ljung-Box test help us in this decision. Note that r_t series is stationary with mean μ_t very close to zero. Using everywhere just r_t instead of innovations e_t but correct is to use innovations/residuals.

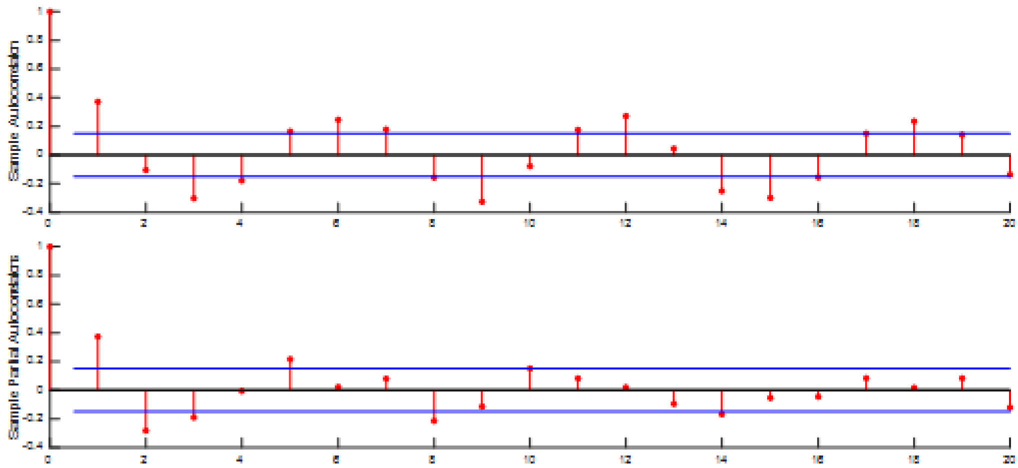

Figure 2: Inflation innovation series
Source: Author's computation

The ACF and PACF show us that food inflation is autocorrelated .We can also reject Ljung-Box test hypothesis with $p-value = 1.5233\text{e-}09$ thus there is at least one non-zero correlation coefficient in $\{\rho(1), \rho(2), \rho(3)\}$ .Next we check for conditional heteroskedasticity of food inflation by examining autocorrelation of squared innovations e_t^2 .

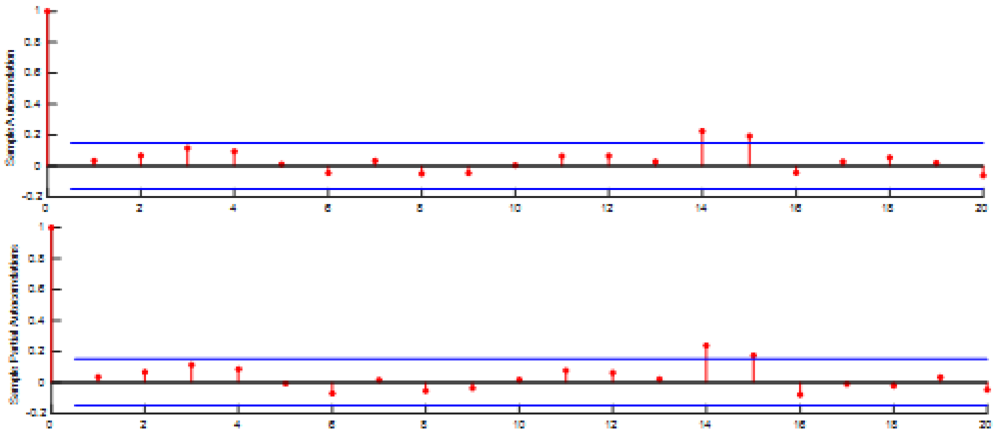

Figure 3: Squared innovation series
Source: Author's computation

The squared innovation series exhibits autocorrelation which tells us that variance of food inflation is significantly autocorrelated thus food inflation is conditionally heteroskedastic.

ARCH test rejects H_0 with $p-value$ =0.0058 in favor of the H_1 hypothesis ,so food inflation innovations e_t are autocorrelated and conditionally heteroskedastic. Let slight modify $r_t = \mu_t + e_t, e_t = \sigma_t \varepsilon_t$ where μ_t is conditional mean and e_t is conditional innovation.

To describe r_t by AR-GARCH models by setting up the ARIMA model objects. This model corresponds to

$$r_t = c + \varphi r_{t-1} + \varphi r_{t-2} + \sigma_t \varepsilon_t \text{ ....(21)}$$

$$\sigma_t^2 = \alpha_0 + \alpha_i e_{t-1}^2 + \beta_i \sigma_{t-1}^2 \text{ ....(22)}$$

Assuming that ε_t is $\varepsilon_t \sim N(0,1)$ or $\varepsilon_t \sim t(v)$ and *i.i.d.* We compare quality of both models using information criterions.

ARIMA(2,0,0) Model (Gaussian Distribution)				
	Value	Standard Error	T Statistic	P Value
Constant	0.60975	0.15799	3.8594	0.00011366
AR{2}	-0.12045	0.077317	-1.5579	0.11925
GARCH(1,1) Conditional Variance Model (Gaussian Distribution)				
	Value	Standard Error	T Statistic	P Value
Constant	0.11388	0.14513	0.78466	0.43265
GARCH{1}	0.92106	0.063158	14.584	3.5762e-48
ARCH{1}	0.044942	0.033943	1.3241	0.18548

Source: Author's estimation

Hence we can rewrite (21) and (22) as

$$r_t = 0.60975 - 0.12045 r_{t-2} + \sigma_t \varepsilon_t \equiv e_t$$

$$\sigma_t^2 = 0.11388 + 0.044942 e_{t-1}^2 + 0.92106 \sigma_{t-1}^2$$

We have just one unknown volatility or conditional variance of inflation σ_t^2 which we can recursively infer. We found out that $\varphi_1 = 0$ in (21) thus has no explanatory power for food inflation as dependent variable. Moreover it seems that innovations autocorrelation is not strong enough to give statistical significance to φ_2 in (21). T-test t-statistic for φ_2 doesn't fall into the critical region so we can't reject hypothesis about zero explanatory power of this coefficient. For innovations from t-distribution we get:

ARIMA(2,0,0) Model (t Distribution)				
	Value	Standard Error	T Statistic	P Value
Constant	0.62394	0.15522	4.0197	5.8268e-05
AR{2}	-0.11344	0.081409	-1.3935	0.16348
DoF	33.857	99.049	0.34182	0.73248
GARCH(1,1) Conditional Variance Model (t Distribution)				
	Value	Standard Error	T Statistic	P Value
Constant	0.10715	0.15353	0.69791	0.48523
GARCH{1}	0.92729	0.067611	13.715	8.2582e-43
ARCH{1}	0.040631	0.035706	1.1379	0.25515
DoF	33.857	99.049	0.34182	0.73248

Source: Author's estimation

Hence we can rewrite (5) and (6) as

$$r_t = 0.62394 - 0.11344 r_{t-2} + \sigma_t \varepsilon_t \equiv e_t$$

$$\sigma_t^2 = 0.10715 + 0.040631 e_{t-1}^2 + 0.92729 \sigma_{t-1}^2$$

So which model choose now? Model Let's examine it quantitatively by AIC, BIC. Before we can compare our models we need to infer log-likelihood objective functions for each of the model. We can also extract final conditional variances volatilities.
AIC = 743.4660 745.2433
BIC = 759.4030 764.3676

So both Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) indicate that AR-GARCH with Gaussian distributed innovations should be chosen.

Now we specify and estimate AR-GJR-GARCH adjusting for asymmetric volatility responses and compare it with better performing AR-GARCH with Gaussian distribution innovations using AIC and BIC. We will define just version with Gaussian distributed innovations.

ARIMA(2,0,0) Model (Gaussian Distribution)				
	Value	Standard Error	T Statistic	P Value
Constant	0.62742	0.15083	4.1599	3.1841e-05
AR{2}	-0.11287	0.081779	-1.3802	0.16753
GARCH(1,1) Conditional Variance Model (Gaussian Distribution)				
	Value	Standard Error	T Statistic	P Value
Constant	0.16162	0.17446	0.92644	0.35422
GARCH{1}	0.90567	0.072424	12.505	7.0076e-36
ARCH{1}	0.095191	0.063572	1.4974	0.13429
Leverage{1}	-0.095191	0.072066	-1.3209	0.18654

Source: Author's estimation

Therefore original AR-GARCH slightly outperforms AR-GJR-GARCH. Actually it is obvious from the output of AR-GJR-GARCH estimate because leverage coefficient is statistically insignificant. Our resulting conditional mean and variance model is AR-GARCH with Gaussian distributed innovations e_t in the following form:

$$r_t = 0.60975 - 0.12045 r_{t-2} + \sigma_t \varepsilon_t \equiv e_t$$

$$\sigma_t^2 = 0.11388 + 0.044942 e_{t-1}^2 + 0.92106 \sigma_{t-1}^2$$

3.4 Simulation in GARCH Model and Conditional Variances

Let's plot food CPI along with AR-GARCH with Gaussian distributed innovations and AR-GARCH with t-distributed innovations.

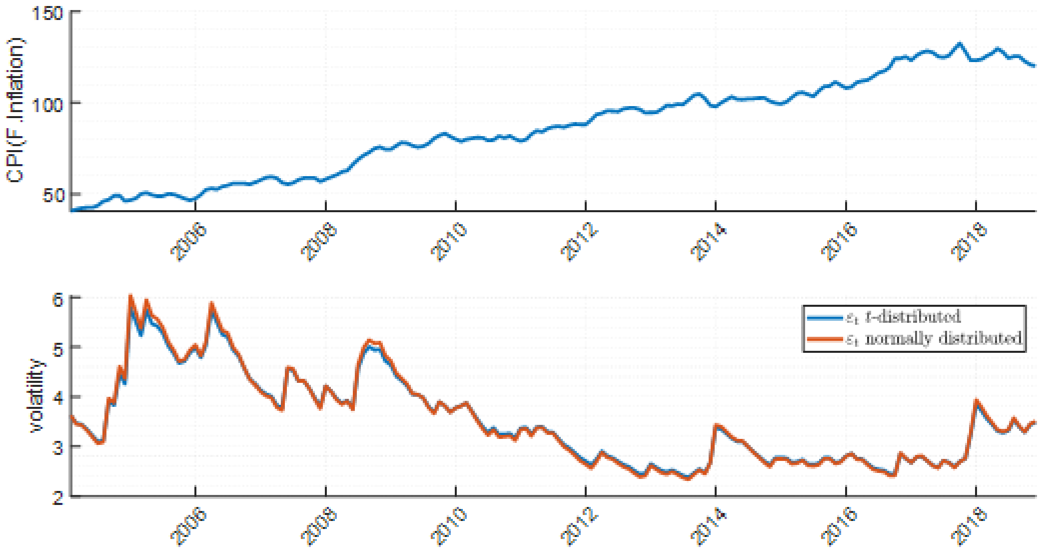

Figure 4: Comparison of generated volatilities
Source: Author's computation

Simulate conditional variance or response paths from a fully specified AR-GARCH model object. That is, simulate from an estimated AR-GARCH model in which specify all parameter values. Plot the average and the 97.5% and 2.5% percentiles of the simulated paths.

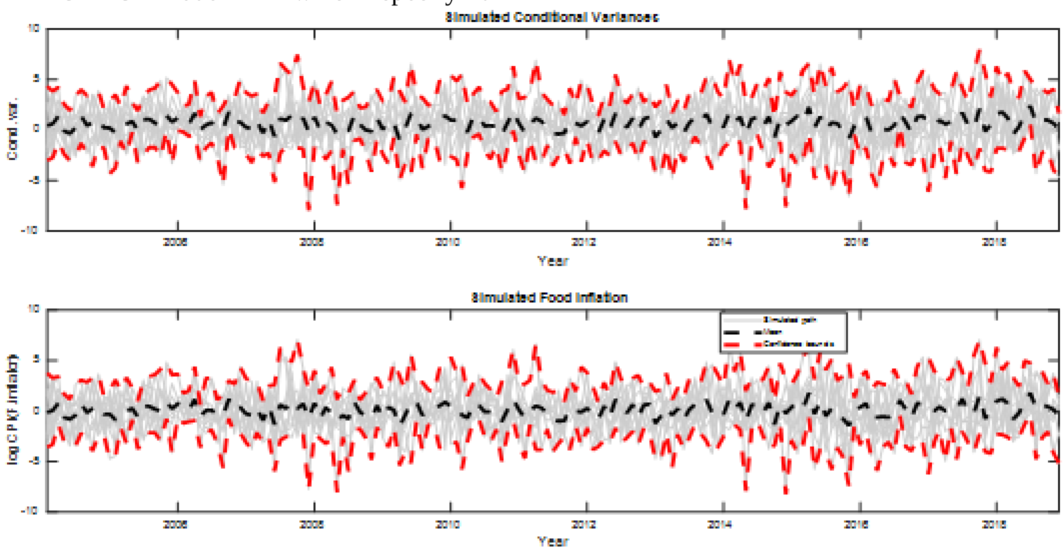

Figure 4: Compare the simulation statistics to the original data
Source: Author's computation

3.5 Forecast of Conditional Variance in GARCH model

Forecast the conditional variance of the food inflation series 6 years into the future using the estimated AR-GARCH model. Specify the entire food inflation series as presample observations. The software infers presample conditional variances using the presample observations and the model.

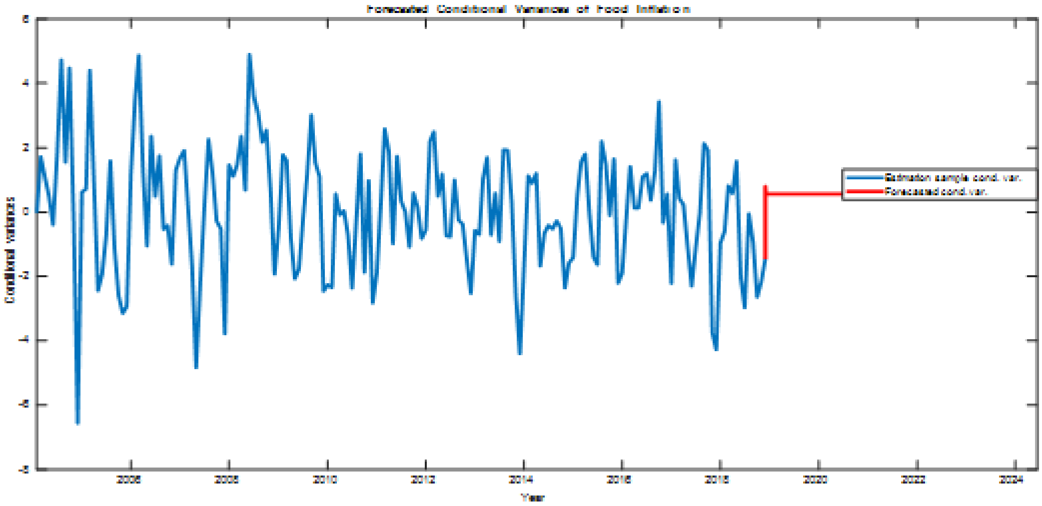

Figure 5: Plot the forecasted conditional variances of the food inflation.
Source: Author's computation

5. Conclusions

This paper introduce GARCH-modelling and its application on Rwanda food inflation data spanning from January, 2004 to December, 2018. On the basis of estimation results of various GARCH Models and diagnostic check has shown that the AR-GARCH with Gaussian distributed innovations is most appropriate specification for modeling food inflation volatility in Rwanda.

The study finds no evidence of asymmetry in the response of food inflation volatility to negative and positive shocks. We further checked the robustness of results in simulation by fitting the food inflation into the estimated AR-GARCH with Gaussian distributed innovations; its density of the fitted Gaussian distribution is the closest to the smoothed estimates of the data and we demonstrate that AR-GARCH with Gaussian with 6-years (72months) forecast from January 2019 to December 2024.

Hence, the AR-GARCH with Gaussian distribution of innovations could be a widely useful tool for modelling the food inflation volatility in Rwanda.

References

[1] Ahiati, V.S. (2007) Discrete Time Series Analysis with ARMA Models. Revised Edition, Holden-Day, Oakland.
[2] Awogbemi, C.A., & Oluwaseyi, A. (2011). Modelling Volatility in Financial Time Series: Evidence from Nigerian inflation rates. Ozean Journal of Applied Sciences, 4 (3), 337 - 350.
[3] Bollerslev, T. (1986). Generalized Autoregressive Cond itional Heteroscedasticity. Journal of Econometrics, 31, 307-327.
[4] Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric Reviews, 11, 143–172.
[5] Bonato, L. 1998. "The Benefits of Price Stability: Some Estimates from New Zealand".Reserve Bank of New Zealand Bulletin Vol. 61(3).
[6] Chatfield, C. (2000) Time Series Forecasting. Chapman and Hall, London.
[7] David, F.H. (2001) Modelling UK Inflation, 1875-1991. Journal of Applied Econometrics, 16, 255-275. http://dx.doi.org/10.1002/jae.615
[8] Engle, R. (1982) The Use of ARCH/GARCH Models in Applied Econometrics. Journal of Economic Perspectives, 15, 157 168.http://dx .doi.org/10.1257/jep.15.4.157
[9] Engle, R. F. (1982).Autoregressive Conditional Heteroscedasticity with Estimate of Variance of United Kingdom inflation. Econometrica, 40, 987-1007.
[10] Hall, R. (1982) Inflation, Causes and Effects. Chicago University Press, Chicago.
[11] Jiang, D. (2011). Inflation and Inflation uncertainty in China. (Unpublished Master's Thesis) School of Economics, The University of Queensland.
[12] King, M. 2005. "Monetary Policy: Practice Ahead of Theory.Bank of England".Mais Lecture of Economics and Finance Vol. 3-4/2008.
[13] Nelson, D. B. (1991). Conditional Heteroscedasticity in Asset returns: A New Approach, Econometrica, 59 , 347 -370.
[14] Ngailo, E.,Luvanda, E. and Massawe, E.S. (2014) Time Series Modelling with Application to Tanzania Inflation Data. Processing, 2, 49-59.
[15] Webster, D. (2000) Webster's New Universal Unabridged Dictionary. Barnes and Noble, New York.