

Application of Autoregressive Conditional Heteroskedasticity (ARCH) Type Models for the Price Volatility of Teff in Ethiopia

Tesfaye Denano^{a*}, Sintayehu Sibera^b

^{a,b}Department of Statistics, Wolaita Sodo University, Ethiopia

^aEmail: tesfaye.d@wsu.edu.et

Abstract

The main aim of this study was to forecast the monthly price volatility of Teff in Ethiopia. The dataset was obtained from the central statistical agency which indexes from January 2014 to June 2021. Thus, ARIMA family models for the mean equation and GARCH family models for the variance equation were employed to forecast the monthly price volatility of Teff. Among the GARCH family models considered in this study, the ARIMA (1,1, 1)-EGARCH (1, 1) model with student t-distributional assumption of residuals was found to be a better fit for the price volatility of Teff. The exchange rate, food inflation rate, non-food inflation rate and fuel oil were found to have a statistically significant effect on the average monthly price volatility of Teff. The asymmetric term was found to be positive and statistically significant in the EGARCH (1, 1) volatility model for Teff. This is an indication that the unanticipated increase in domestic prices had a larger impact on domestic price volatility than the unanticipated decrease in the domestic price of Teff. Finally, various forecast accuracy measurement statistics indicate that the estimated ARIMA (1,1,1) model is good enough to describe the domestic price of Teff. Moreover, the out-of-sample forecast indicates that the domestic price of Teff is increasing. The in-sample forecast using the best-fit asymmetric model, which is the EGARCH (1, 1) model, indicates that the domestic price volatility of Teff remained at almost a constant level around the beginning and end of the study. But it increased steadily at an increasing rate from December 2019 to January 2021, then dropped sharply at an increasing rate till March 2021.

Keywords: Monthly price volatility of Teff; ARIMA; ARCH Type Models; Forecasting; Ethiopia.

* Corresponding author.

1. Introduction

Natural and man-made disasters, such as earthquakes, war, and the Coronavirus (COVID 19), are extremely likely to price volatility of Teff in Ethiopia. Disruptions in the supply chain are defined as high-impact, low-frequency occurrences that alter the supply chain's structural architecture and have a substantial influence on performance [1, 7]. Ripple effects are the consequences of a disturbance propagating across a supply chain. COVID 19 will have short to long-term economic consequences in Ethiopia. Forecasting agricultural commodity prices is critical for policymakers as well as different stakeholders in the agricultural commodity marketing chain, from farmers to consumers [9,10, 16, 17]. The majority of agricultural price series may be modeled as time series data, in which data is gathered at regular intervals across time. The Box Jenkins' autoregressive integrated moving average (ARIMA) approach dominated time-series forecasting for a long time, until the necessity to cope with volatile data became apparent. Surprisingly, many agricultural commodity price statistics are noisy and erratic by nature. This is due to the fact that agricultural commodity prices react quickly to actual and perceived changes in supply and demand circumstances, and weather-related fluctuations in farm productivity exacerbate the problem. Asymmetric occurrences may also occur in price series, which tend to behave differently when the economy enters or exits a recession. Volatility is a rapid and unexpected increase or decrease in a series that can irritate investors. It is a well-known fact that price volatility may destabilize farm revenue and prevent farmers from making optimal investments and resource use [15]. This has the potential to divert much-needed resources away from agriculture. When a few error-terms are bigger than the rest and are responsible for the series' distinctive behavior, the series is said to be volatile; this phenomenon is known as heteroskedasticity. Reference [4, 5] extended the model as the Generalized ARCH (GARCH) model for parsimonious representation of ARCH. The conditional variance in the GARCH model is likewise a linear function of its own lags. This model, like ARCH, is a weighted average of past squared residuals, but it has decreasing weights that never reach zero. The GARCH models are frequently utilized for economic and financial series modeling and forecasting. Further progress was achieved by combining the AR specification with the ARCH/GARCH models, which proved to be more effective in forecasting volatility. Non-linear models have been widely employed by numerous academics for more than two decades, who have discovered diverse combinations of AR-GARCH models that are suited for varied scenarios [12]. In time-series forecasting, however, the GARCH (1, 1) model is the most commonly used GARCH formulation. Various time-series have revealed both symmetric and asymmetric trends in recent years. The GARCH model has been shown to be somewhat inefficient in modeling and forecasting such series due to its nature of dealing only with size rather than the positivity or negativity of the shocks. As a result, the need for expansion of the GARCH family model was felt, and Nelson was the first to respond. Therefore, this study is motivated to forecast prices volatility of Teff in Ethiopia.

2. Methods and Materials

2.1. Sources of Dataset

Data for the study were obtained from Central statistical agency (CSA) and national bank of Ethiopia (NBE). Period from Jan 2014 to June 2021 were used for estimation process. The price volatility of Teff, exchange rate and fuel oil price data are obtained from NBE and general food inflation and general non-food inflation data are

obtained from CSA. The variables of interest in this study are price volatility of Teff which is to be used as a dependent variable, and fuel oil price, exchange rate general food inflation and general non-food inflation are exogenous variables used to model and forecast the volatility of the price of Teff (general price inflation) in Ethiopia.

2.2. Model Specification

Time series models are mainly concerned with analysis (estimation) and forecasting of different equations concerning stochastic components since time series is a sequence of real random variables (y_t) defined on the same probability space. Moreover, it is an example of a discrete-time stochastic process [14]. In this study, some of the time series models such as conditional mean and conditional variance (GARCH family) models are applied to model and forecast price inflation volatility in Ethiopia.

2.2.1. Stationarity and Unit-Root Problem

Generally the concept of stationarity can be summarized by the following conditions. A time series y_t is said to be stationary if:

$$\text{mean} = E(y_t) = E(y_{t-s}) = \mu \quad 1$$

$$\text{variance} = E(y_t - \mu)^2 = E(y_{t-s} - \mu)^2 = \delta^2_y \quad 2$$

$$\text{Covariance} = E(y_t - \mu)(y_{t-s} - \mu) = E(y_{t-j} - \mu)(y_{t-j-s} - \mu) = \gamma_s \quad 3$$

Where μ , δ^2_y & γ_s are all time invariant

The assumption of stationarity is somewhat unrealistic for most macroeconomic variables. A non-stationary process arises when at least one of the conditions for stationarity does not hold. Let us consider an autoregressive process of order one (AR (1) process):

$$y_t = \rho y_{t-1} + \varepsilon_t \quad 4$$

Where ε_t denotes a serially uncorrelated white noise error term with a mean of zero and a constant variance. Non-stationarity can originate from various sources but the most important one is the presence of so-called “unit roots”. Equation (4) is said to be a unit root process when $\rho = 1$

Let p_t , $t = 1, 2, 3, \dots$ be the price of a commodity at time period t (t in days, months, etc). Instead of analyzing p_t , which often displays unit-root behavior and thus cannot be modeled as stationary, we often analyze log returns on p_t [11]:

$$Y_t = \left(\frac{p_t}{p_{t-1}} \right) = \log \left(1 + \frac{p_t - p_{t-1}}{p_{t-1}} \right) \quad 5$$

The series Y_t log- return series, displays many of the typical characteristics in financial time series such as volatility, clustering and leptokurtosis.

2.2.2. The Mean Models

2.2.2.1. The Autoregressive Moving Average (ARMA)

In general, an ARMA model is denoted by ARMA (p, q), where p and q are the orders of autoregressive and moving average components, respectively.

ARMA (p, q) mean model [5] is given by:

$$Y_t = K + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \dots + \varphi_q \varepsilon_t \quad 6$$

Where K is a constant and ε_{t-k} ($k=1, 2 \dots q$) are the values of the previous random shock.

2.2.2.2. The Autoregressive Integrated Moving Average (ARIMA) Model

The ARIMA model was introduced by Box and Jenkins (hence also known as Box-Jenkins model) in 1960s for forecasting a variable. ARIMA models consist of unit-root non-stationary time series which can be made stationary by the order of integration d . The general form of ARIMA (p, d, q) is written as:

$$\Delta^d \varpi_p(B) Y_t = \phi_0 + \eta_q(B) \varepsilon_t \quad 7$$

Where $\varpi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, $\eta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$
 d is the order of integration and B is the Backward shift operator

2.2.3. The Variance Models

The Box-Jenkins time series model such as Autoregressive (AR), Moving Average (MA) and ARMA are often very useful in modeling general time series data. However, they all require the assumption of homoscedasticity (or constant variance) for the error term in the model. But this may not be appropriate when dealing with some special characteristics of financial and agricultural price time series. This led to the introduction of Autoregressive Conditional Heteroskedasticity (ARCH) model which was proposed by [8]. After identifying the presence of ARCH effects, separate GARCH, TGARCH and EGARCH models have been employed in this study to investigate the pattern of price volatility and its determinants for cereal crops (wheat and barley) and pulse seeds (pea and bean) with joint estimation of a mean and a conditional variance equation.

2.2.3.1. The Autoregressive Conditional Heteroskedasticity (ARCH) model

The ARCH model for the variance of the errors, denoted by ARCH (Q), was proposed by Engle (1982). The conditional variance is given by:

$$\eta_t = \delta_t u_t$$

$$\delta_t^2 = \beta_o + \beta_1 \varepsilon_{t-1}^2 + \dots + \beta_p \varepsilon_{t-p}^2$$

$$\delta_t^2 = \beta_o + \sum_{j=1}^p \beta_j \varepsilon_{t-j}^2$$

Where, u_t is IID normal residual with mean zero and unit variance and δ_t^2 is the conditional variance of the residuals at time t, i.e., $\text{var}(\eta_t | \eta_{t-1}, \eta_{t-2}, \dots) = \delta_t^2$. This indicates that the current value of the variance of the errors possibly depend upon previous squared error terms. We impose the non-negativity constraints $\beta_o, \beta_j > 0, j = 1, 2, 3, \dots, p$

$$\eta_t = \delta_t u_t \quad 8$$

$$\delta_t^2 = \beta_o + \beta_1 \varepsilon_{t-1}^2 + \dots + \beta_p \varepsilon_{t-p}^2 + \varphi_t \delta_{t-1}^2 + \dots + \varphi_q \delta_{t-q}^2 \quad 9$$

$$\delta_t^2 = \beta_o + \sum_{j=1}^p \beta_j \varepsilon_{t-j}^2 - \sum_{k=1}^q \varphi_k \delta_{t-k}^2 + \varepsilon_t \quad 10$$

$$\text{restrictions } \beta_o > 0, \beta_p \geq 0, \varphi_q > 0$$

2.2.3.2. The Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) model

In order to capture possible asymmetry exhibited by financial time series, a new class of models, termed the asymmetric ARCH models, was introduced. The most popular model proposed to capture the asymmetric effects is exponential GARCH, or EGARCH model. The ARMA(p, q)-EGARCH (P, Q) model is given as:

$$y_t = \gamma_0 + \sum_{i=1}^p \gamma_i y_{t-i} - \sum_{j=1}^q \varpi_j \varepsilon_{t-j} + \varepsilon_t \quad 11$$

$$\ln(\delta_t^2) = \alpha_o + \sum_{i=1}^q \alpha_i \left| \frac{\varepsilon_{t-i}}{\delta_{t-i}} \right| + \sum_{i=1}^R \lambda_i \left(\frac{\varepsilon_{t-i}}{\delta_{t-i}} \right) + \sum_{j=1}^q \beta_j \ln(\delta_{t-j}^2) \quad 12$$

In contrast to the GARCH model, no restrictions need to be imposed on the model parameters since log transformed conditional variance overcomes the positivity constraint of coefficients in EGARCH models. Note that the left hand side is the log of the conditional variance. This implies that the leverage effect is exponential, rather than quadratic and that forecasts of the conditional variance are guaranteed to be non-negative. In this model specification, $\beta_1, \beta_2, \dots, \beta_p$ are the GARCH parameters that measure the impact of past volatility on the current volatility.

2.2.3.3. The Threshold Generalized Autoregressive Conditional Heteroskedasticity (TGARCH) model

The number of possible conditional volatility formulations is vast. The threshold GARCH, or TGARCH (P,Q), model is one of the widely used models introduced by Zakeian (1990) and Glosten and his colleagues (1993). The TGARCH model with mean and conditional variance equations is given as:

$$y_t = \gamma_0 + \sum_{i=1}^p \gamma_i y_{t-i} - \sum_{j=1}^q \varpi_j \varepsilon_{t-j} + \varepsilon_t \quad 13$$

$$\delta_t^2 = \eta_0 + \sum_{j=1}^p \eta_j \varepsilon_{t-j}^2 + \sum_{k=1}^q \varphi_k m_{t-k} + \sum_{j=1}^q \beta_j \delta_{t-j}^2 + \varepsilon_t \quad 14$$

where $m_{t-i} = 1$ if $\varepsilon_{t-i} \geq 0$ and $m_{t-i} = 0$ otherwise in (equation 13). The TGARCH model allows a response of volatility to news with different coefficients for good and bad news. That is, depending on whether ε_{t-i} is above or below the threshold value of zero, ε_{t-i} has different effects on the conditional variance, δ_t^2 : when ε_{t-i} is positive, the total effects are given by $\eta_i \varepsilon_{t-i}$ when ε_{t-i} is negative, the total effects are given by $(\eta_i + \lambda_i) \varepsilon_{t-i}$. So, one would expect λ_i to be positive for bad news to have larger impacts. The presence of leverage effects can be tested by the hypothesis that $\lambda_i = 0$. The impact is asymmetric if $\lambda_i \neq 0$. In this study, food inflation rate, non-food inflation rate, fuel oil price and exchange rate were introduced into the conditional variance equation as independent variables in order to determine the impact of these variables on the volatility of average monthly price returns under consideration. The conditional variance equation of GARCH (P,Q) with explanatory variables for each cereal crops and pulse seeds is given by:

$$\delta_t^2 = \eta_0 + \sum \eta_i \varepsilon_{t-i}^2 + \sum \beta_j \delta_{t-j}^2 + \psi' X_t$$

Where, $X_t = (x_{1t}, \dots, x_{kt})$ is a vector of explanatory variables and $\psi = (\psi_1, \dots, \psi_k)'$ is a vector of regression coefficients of the explanatory variables. Assuming the presence of asymmetric effect on the GARCH family model, the conditional variance equations for EGARCH (P,Q) and TGARCH(P,Q) with explanatory variables are given by:

$$\ln(\delta_t^2) = \alpha_0 + \sum_{i=1}^q \alpha_i \left| \frac{\varepsilon_{t-i}}{\delta_{t-i}} \right| + \sum_{i=1}^R \lambda_i \left(\frac{\varepsilon_{t-i}}{\delta_{t-i}} \right) + \sum_{j=1}^q \beta_j \ln(\delta_{t-j}^2) + \psi' X_t$$

$$\delta_t^2 = \eta_0 + \sum_{j=1}^p \eta_j \varepsilon_{t-j}^2 + \sum_{k=1}^q \varphi_k m_{t-k} + \sum_{j=1}^q \beta_j \delta_{t-j}^2 + \psi' X_t \quad 15$$

Assumptions of the Models

- i. The expected value of the error term is zero, i.e. $E(\varepsilon_t) = 0$.
- ii. The variance of the error terms is conditionally heteroskedastic.

- iii. Error terms are independent having normal or student-t or GED distribution with mean zero and variance δ_t^2
- iv. There is no serial autocorrelation among successive error terms.
- v. No severe Multicollinearity exists among explanatory variables.

2.3. Procedures for Model Building

2.3.1. Testing for the Presence of Unit Root

The widely used unit-root tests are Augmented Dickey Fuller test [6, 18]. Once the presence of unit root (non-stationarity) is confirmed, the series needs to be differenced to achieve stationarity.

2.3.1.1. The Dickey Fuller Test

The simplest form of the Dickey-Fuller [6] test amounts to estimating

$$y_t = \rho y_{t-1} + \varepsilon_t \quad (16)$$

One could use a t- test to test the hypothesis $\rho = 1$ (unit root) against $\rho < 1$ alternatively, one can rearrange the model as follows:

$$\Delta y_t = (\rho - 1)y_t + \varepsilon_t \quad (17)$$

The simple Dickey-Fuller unit root test described above is valid only if the series is an AR (1) process. If the series is correlated at higher order lags, the assumption of white noise disturbances is violated. The Augmented Dickey-Fuller test constructs a parametric correction for higher order correlation by assuming that the series follows an AR (p) process and adding lagged difference terms of the dependent variable to the right-hand side of the test regression. The ADF test is comparable with the simple DF test, but is augmented by adding lagged values of the first difference of the dependent variable as additional regressors which are required to account for possible occurrence of autocorrelation. Consider the AR (p) model:

$$\Delta y_t = \mu + \beta y_{t-1} + \sum_{i=1}^p \phi_i \Delta y_{t-i} + \varepsilon_t \quad (18)$$

The presence of a unit root is tested in a similar manner as described above. If the null hypothesis $H_0 : \beta = 0$ is not rejected, then we need to difference the data to make it stationary or we need to put a time trend in the regression model to correct for the variables' deterministic trend.

2.3.1.2. The Phillips and Perron Test

An important assumption of the DF test is that the error terms ε_t are independently and identically distributed.

The ADF test adjusts the DF test to take care of possible serial correlation in the error terms by adding lagged difference terms of the dependent variable. Phillips and Perron use nonparametric statistical methods to take care of the serial correlation in the error terms without adding lagged difference terms. Examining AC patterns within a time series is an important step in many statistical analyses. The autocorrelation coefficient is the correlation between y_t and y_{t-k} (separated by k periods apart) and the formal expression at time lag k is:

$$\rho_k = \frac{\text{cov}(y_t, y_{t-k})}{\sqrt{\text{var}(y_t) \text{var}(y_{t-k})}} \quad (19)$$

Where, $\text{cov}(y_t, y_{t-k})$ is the auto-covariance between y_t and y_{t-k} for $k = 0, 1, 2 \dots$. The ACF is a listing or graph of the sample autocorrelations at lag $k = 0, 1, 2 \dots$. The partial autocorrelation at lag k is the correlation between y_t and y_{t-k} after removing the effect of $y_{t-1}, \dots, y_{t-k+1}$ and the formal expression is given as:

$$\gamma_k = \frac{\text{cov}(y_t, y_{t-k} / y_{t-1}, \dots, y_{t-k+1})}{\sqrt{\text{var}(y_t / y_{t-1}) \text{var}(y_{t-k} / y_{t-k+1} + 1)}} \quad (20)$$

Where $\text{cov}(y_t, y_{t-k} / y_{t-1}, \dots, y_{t-k+1})$ is the auto covariance between y_t and y_{t-k} after removing the effect of $y_{t-1}, \dots, y_{t-k+1}$ for $k = 0, 1, 2 \dots$. However, for GARCH family models the autocorrelation and partial autocorrelation coefficients are computed

$$\rho_k = \frac{\text{cov}(y_t^2, y_{t-k}^2)}{\sqrt{\text{var}(y_t^2) \text{var}(y_{t-k}^2)}} \quad (21)$$

Where $\text{cov}(y_t^2, y_{t-k}^2)$ is the auto covariance between y_t^2 and y_{t-k}^2 for $k = 0, 1, 2 \dots$. The partial autocorrelation coefficient at lag k can be computed as:

$$\gamma_k = \frac{\text{cov}(y_t^2, y_{t-k}^2 / y_{t-1}^2, \dots, y_{t-k+1}^2)}{\text{var}(y_t^2 / y_{t-1}^2, \dots, y_{t-k+1}^2)} \quad (22)$$

Where $\text{cov}(y_t^2, y_{t-k}^2 / y_{t-1}^2, \dots, y_{t-k+1}^2)$ is the auto covariance between y_t^2 and y_{t-k}^2 after removing the effect of $y_{t-1}^2, \dots, y_{t-k+1}^2$ for $k = 0, 1, 2, \dots$. The plot of the autocorrelation and partial autocorrelation functions of the ordinary and standardized residuals at time lag $k = 0, 1, 2 \dots$ can be used for checking the presence of

ARCH effects in the residuals from mean equation. In case of ARCH effects, ACFs of squared residuals should dies down slowly (i.e. ACFs of squared residuals should not be white noise). The ACFs are also used for checking model adequacy in fitted GARCH (p, q) family process, i.e. the ACFs of standardized residuals should be indicative of a white noise process if the model is adequate.

2.3.2. Testing ARCH Effects

The presence of ARCH effect (whether or not volatility varies over time) has to be tested through the squared residuals of the series[20]. According to [20], there are two available methods to test for ARCH effects.

2.3.2.1. Ljung-Box Test

It was developed by [6] and modified by Ljung and Box (1978) and tests the joint significances of serial correlation in the standardized and squared standardized residuals for the first k lags instead of testing individual significance. They suggested testing the hypothesis:

$$H_o : \rho_1 = \rho_2 = \dots = \rho_k = 0 \text{ (the ACF for the first k lags of the squared residuals series is zero)}$$

$$H_a : \text{at least one } \rho \text{'s different}$$

They suggested the statistic:
$$Q(k) = n(n-2) \sum_{j=1}^k \frac{d_j^2}{n-j}$$

Where n denotes the length of the series after any differencing and dj denotes the squared residual from equation (2). They showed that under the null hypothesis Q(k) is asymptotically distributed as chi-square with (k-p-q) degree of freedom, where k is the maximum lag considered, p and q are the order of the AR and MA from equation (2), respectively.

2.3.2.2. Lagrange Multiplier (LM) Test

This test was suggested by [8] and used to test the significance of serial correlation in the squared residuals for the first q lags. The steps to derive the test statistic for LM test are:

- ✚ Estimate the mean equation ε , ε^*
- ✚ Obtain the residuals
- ✚ Then regress current squared residual on lagged squared residuals and a constant

$$\hat{\varepsilon}_t^2 = \gamma_o + \gamma_1 \hat{\varepsilon}_{t-1}^2 + \dots + \gamma_q \hat{\varepsilon}_{t-q}^2$$

The null hypothesis is that, $\gamma_o = \gamma_1 = \dots = \gamma_q = 0$

The test statistic $n \cdot R^2$ is distributed as chi-square with q degrees of freedom, where R^2 is the coefficient of determination from equation (15) and n is number of observations. The rejection of the null hypothesis indicates the presence of ARCH effects. It is important to apply the LM test on the residuals from the mean equation ARMA model (not GARCH model). Also one does not test directly for GARCH effects; if ARCH effect exists, GARCH model can also be considered.

Test of Normality

When dealing with GARCH family models, the data is first tested for normality (i.e. whether the returns follow a normal distribution). In statistics, the JB test is a test of departure from normality based on the sample kurtosis and Skweness. The null hypothesis states that the observations come from a normal distribution. The test statistic is:

$$JB = \frac{n}{6} * \left(S^2 + \frac{(k-3)^2}{4} \right) \quad (23)$$

Where n is the number of observations, S is the sample Skweness and K is the sample kurtosis. Under the null hypothesis, the Jarque-Bera statistic is distributed as chi-square distribution with two degrees of freedom.

2.3.3. Model Order Selection in GARCH Family Model

For checking model adequacy in fitted GARCH (p, q) family process, The ACFs can be used i.e. the ACFs of standardized residuals should behave as a white noise process if the model is adequate for the data. Since GARCH models can be treated as ARMA models for squared residuals, traditional model selection criteria such as the Akaike information criterion and the Schwartz Bayesian information criterion (SBIC). After selecting appropriate lag order, the final models can be selected using information criteria's or penalty function statistics such as Akaike Information Criterion or Bayesian Information Criterion. The AIC and BIC are a measure of the goodness of fit of an estimated statistical model. After ranking competing models according to their AIC or BIC, the one having the lowest information criterion value can be selected as the best model among several models with similar distributional assumptions. These information criterion measures a model by how close its fitted values tend to be to the true values, in terms of a certain expected value. The information criterion finds the model that best fits the data with a minimum of free parameters but also includes a penalty that is an increasing function of the number of estimated parameters. Moreover, forecast accuracy measure statistics among the competing candidate models could also help in making a decision. Minimum of free parameters but also includes a penalty that is an increasing function of the number of estimated parameters. This penalty discourages over fitting. In the general case, the AIC and BIC take the form as shown below [20]. The formal expressions for the above criteria in terms of the log- likelihood are:

$$AIC = -2\ln(L) + 2k \quad (24)$$

$$BIC = -2\ln(L) + k\ln(n) \quad (25)$$

Where n = number of observations

K = number of parameters estimated

L = value of the likelihood function $\log(L(\delta^2))$

2.3.4. Model Adequacy Checking

After a GARCH family model has been fit to the data, the adequacy of the fit has been evaluated using a number of graphical and statistical diagnostics.

3. Results and Discussion

The data in this study consist of monthly price volatility of Teff, monthly food consumer price index, non-food consumer price index, monthly fuel oil price (in US dollar) and monthly exchange rate (in birr against US dollar) in Ethiopia for the period spanning from January 2014 to June 2021

3.1. Test for Stationarity

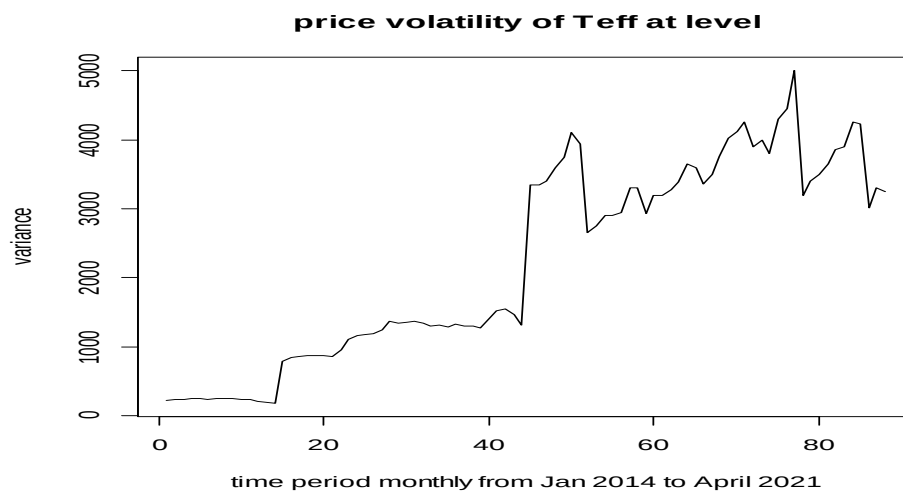


Figure 1: Time plot of monthly price of Teff at level

Source: author's estimation using Eviews-8

The time series under consideration should be checked for Stationarity before one attempts to fit a suitable

model. The Stationarity of each series can be tested using the Augmented Dickey-Fuller test and the Phillips-Perron test. The results of ADF and PP tests with intercept and trend both at level and at first difference for the series are presented in Tables 2. Test results presented in Table 2 indicate that the null hypothesis that the series at levels contain unit root could not be rejected for the price volatility of Teff. The graphical description of the trend of average monthly domestic price of Teff and its returns are as shown in figure-1 below. From the figure it can be observed that average monthly domestic price of Teff shows increasing pattern followed by a period where price is decreased up to the end of the forecasting periods under consideration.

The absolute return series were constructed for monthly price of Teff in order to examine the presence of volatility in the series. As can be seen from Figure-2, periods of high volatility is followed by periods of high volatility and periods of low volatility is followed by periods of low volatility for the series.

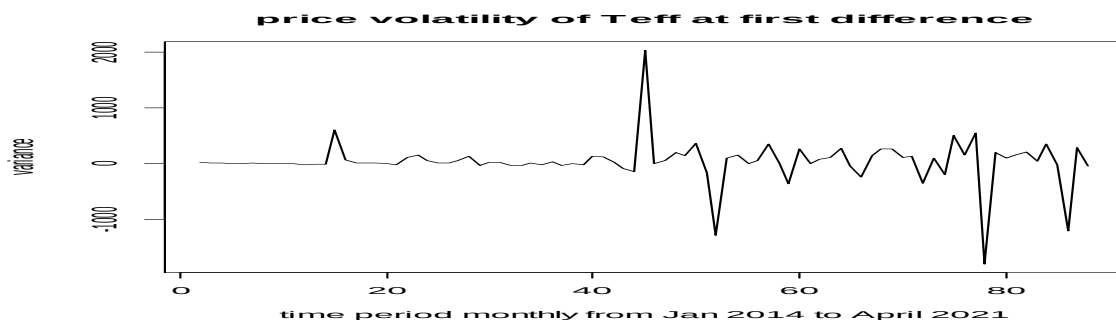


Figure 2: Time plot of monthly price of Teff at first difference

If a time series data is non-stationary, it is necessary to look for possible transformations that might induce Stationarity. In practice, researchers usually transform financial data series into return forms. Table 1 summarizes the unit root test of the return series for the price volatility of Teff. The table shows that the null hypothesis of unit root would be rejected. Hence the return series of the price volatility of Teff are stationary at first difference.

Table 1: ADF and PP Unit root test result the return series of the price volatility of Teff

Phillips-Perron Test			Augmented Dickey-Fuller Test		Order of Identification
Variables	intercept , with Trend		intercept , with Trend		
	At level	First difference	At level	First difference	
Price of Teff	-2.0687 (0.1204)	-2.249 (0.000)*	-4.185 (0.216)	-8.494 (0.000)*	I(1)
Exchange Rate	-12.9555 (0.120)	-11.835 (0.000)	-13.231 (0.100)	-10.735 (0.000)	I(1)
Fuel Oil Price	-12.548 (0.231)	-12.541 (0.000)	-11.641 (0.211)	-9.641 (0.000)*	I(1)
Non-Food	-8.301 (0.291)	-9.411 (0.000)*	-10.411 (0.000)*	-10.358 (0.000)*	I(1)
Food (stable)	-8.1254 (0.180)	-7.625 (0.000)*	-8.725 (0.000)	-8.6255 (0.000)	I(1)

3.2. Specification Conditional Mean Equation

In order to model the volatility of the return series, first we need to specify their conditional mean equation, since the return for current time will depend on returns in previous periods (autoregressive component) and the error terms in current and previous periods (moving average component). To model the conditional mean equation for the series the Box-Jenkins modeling approach has been used. The Box-Jenkins modeling approach of a stationary time series involves the following four steps: model identification, model estimation, diagnostic checking and forecasting. We have seen that the return of domestic price of Teff becomes a stationary time series at level. So, the model that we are looking at is ARMA (p, q). Hence we have to identify a good-fit model, estimate its parameters, apply diagnostic checking for the residuals and finally achieve our objectives of in-sample and out-of-sample forecasting of the domestic price of Teff. In the specification of the mean equation, the sample ACF and PACF plots of the stationary series can be used to tentatively identify the order of autoregressive terms and/or moving average terms. To specify the conditional mean equation for the series, comparison of various AR (p), MA (q) and ARIMA (p, d, q) models are performed and the one with smallest information criteria is selected. In this study, AR (0-3) and MA (0-3) were considered since the return series show insignificant spikes for all of the lags. Among the various ARIMA models considered, ARIMA (1, 1, 1) model possesses minimum AIC and BIC and exhibits no serial correlation. Therefore, ARIMA (1, 1, 1) model is the best-fit model for the conditional mean equation for the return series of the price volatility of Teff. The maximum likelihood estimation method for monthly return series of price volatility of Teff are summarized in Table 2 below.

Table 2: Parameter estimate of ARIMA (1, 1, 1) model

Series	Coefficient	Std.error	t-statistic	p-value
Constant	17.98740	11.28299	1.594206	0.0147**
AR(1)	0.697393	0.080590	8.653640	0.0000*
MA(1)	-1.159182	0.070436	-16.45719	0.0000*

Note: * and ** indicates significant at 1% and 5% level, respectively

From Table 2, the estimated parameters of the moving average term and the constant term are statistically significant at 5% level of significant. So the equation of the fitted ARIMA model is given by:

$$y_t = 17.9874 + 0.697393y_{t-1} - 1.159182\varepsilon_{t-1} + \varepsilon_t$$

3.3. Diagnostic checking

Before we consider the fitted model as a better fit and interpret its findings, it is essential to check whether the model is correctly specified, that is, whether the model assumptions are supported by the data. If some key model assumptions seem to be violated, then a new model should be specified until it provides an adequate fit to the data. The presence of serial correlation in the residuals

was tested using the Lagrange Multiplier (LM) and Ljung-Box tests for ARIMA(1,1,1) model. The null hypothesis asserts that there is no serial correlation in the residuals up to lag 24. The results of this examination are summarized in Table-3. The Breusch–Godfrey serial correlation LM test results in Table 5 provide an evidence of absence of serial correlation in the residuals of the mean equation, since the null hypothesis is not rejected at 1% level of significance. The Ljung-Box test (Table 5) also indicates that there is no significant serial correlation for lags up to 24 for the series. Hence, we can conclude that there is no significant serial correlation in the residuals of the mean equation.

Table 3: Serial Correlation Test

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. **	. **	1	0.263	0.263	6.1657	0.113
. * .	. .	2	0.096	0.029	6.9931	0.130
. .	. .	3	0.005	-0.029	6.9958	0.172
. .	. .	4	-0.049	-0.049	7.2130	0.125
. .	. .	5	0.009	0.039	7.2208	0.205
. .	. .	6	0.020	0.016	7.2569	0.298
. * .	. * .	7	0.206	0.208	11.305	0.126
. * .	. * .	8	0.176	0.079	14.318	0.174
. .	. .	9	0.045	-0.046	14.518	0.105
. * .	. * .	10	-0.074	-0.102	15.061	0.130
. .	. .	11	-0.050	0.013	15.316	0.168
. .	. .	12	-0.052	-0.025	15.592	0.211
. * .	. .	13	-0.074	-0.060	16.158	0.241
. .	. .	14	-0.031	-0.045	16.259	0.298
. * .	. * .	15	-0.073	-0.111	16.835	0.329
. * .	. .	16	-0.075	-0.061	17.447	0.357
. * .	. .	17	-0.083	-0.013	18.197	0.377
. * .	. .	18	-0.078	-0.018	18.878	0.399
. .	. .	19	-0.057	-0.026	19.249	0.441
. * .	. .	20	-0.086	-0.053	20.092	0.452
. * .	. .	21	-0.077	-0.030	20.782	0.472
. * .	. .	22	-0.097	-0.040	21.884	0.467
. * .	. .	23	-0.070	0.002	22.481	0.491
. * .	. .	24	-0.080	-0.037	23.255	0.505

Source: author's estimation using Eviews-8

3.4. Test of normality of the residuals

To investigate whether the residuals of the fitted model in the mean equation are normally distributed or not, the Jarque-Bera test has been applied. The results are reported in table 4:

Table 4: Normality test of residuals from the mean equation

Variable	Skweness	Kurtosis	Jarque-Bera Statistic	p-value
Price return series	-0.7011	7.4585	30.351	0.3115

Source: author's estimation using Eviews-8

As we can see from Table 4 the Jarque-Bera statistic is insignificant and hence, there is no a significant evidence to reject the null hypothesis of normality. This indicates that the residuals of the fitted model are approximately normally distributed for the series under consideration.

3.5. Forecasting of ARIMA (1, 1, 1) model

One of the fundamental applications of time series analysis or developing a time series model is forecasting. The previous discussion has shown that the ARIMA (1, 1, 1) model is a good-fit model to describe the monthly price of Teff. In this section we examine the forecasting accuracy of the fitted model and then make in-sample and out-of-sample forecast.

3.5.1. Evaluation of in-sample forecast

The forecasting performance of a model can be examined by the standardized statistical tools such as root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and Theil's inequality.

Table 5: In-sample forecast accuracy measures for ARIMA (1, 1, 1) model

Accuracy measures	Values
Root Mean Squared Error	0.62381
Mean Absolute Error	0.041512
Mean Absolute Percentage Error	361.1380
Theil Inequality Coefficient	0.031345
Bias Proportion	0.00112
Variance Proportion	0.25431
Covariance Proportion	0.241941

Here as we have seen from Table 5 most of the statistics are small. Moreover, the U statistic is 0.031345, which is approximated to zero. Besides that the bias and the variance proportion are also close to zero. This indicates that the forecasting performance of the model is good.

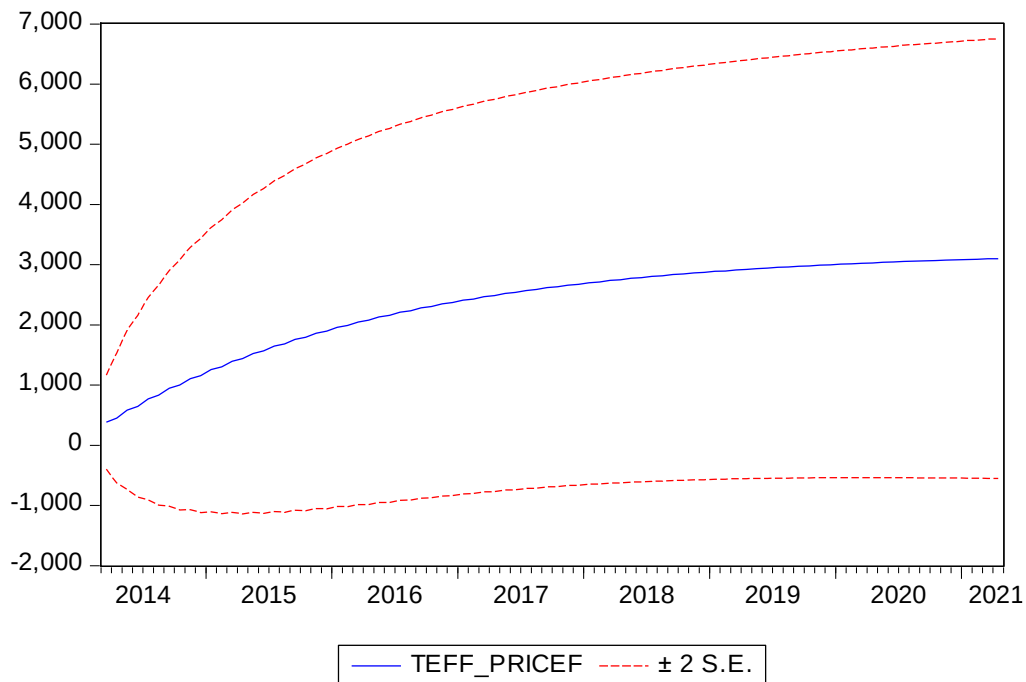


Figure 3: Out-of-sample forecast for ARIMA (1, 1, 1) model

Source: author's estimation using Eviews-8

As we can see from the figure-3 above the forecasting trend of monthly price of Teff shows an increasing pattern over the forecasting periods under consideration

3.6. Test for ARCH effects

To proceed with volatility modeling, ARCH effects (whether or not volatility varies over time) in the residuals from the selected ARMA model should be tested. The approaches for testing heteroskedasticity (ARCH effect) were the Lagrange Multiplier (LM) test proposed by [5, 8] test for squared residuals. To proceed with volatility modeling ARCH effects (whether or not volatility varies over time) in the residuals from the selected ARIMA (1, 1, 1) model should be tested. The confirmation of the presence of ARCH effect indicates that the volatility in the average monthly return price volatility of Teff is time varying and appropriateness of employing GARCH family model (table-6). These results indicate that the respective return series under consideration have a non-constant variance (heteroskedasticity) and need to be modeled using GARCH family models.

Table 6: ARCH effect test using LM test for squared residuals of the fitted model (Heteroskedasticity Test: ARCH)

F-statistics	1.314545	Probability	0.00235
Obs*R-squared	8.84546	Probability	0.002512

Source: author's estimation using Eviews-8

3.7. Volatility Models

3.7.1. GARCH Family Model

Once the ARCH effects are determined, then the optimal lag specifications for GARCH family models were determined prior to the construction of the final model to investigate the determinants of domestic price volatility. Even though there is consensus that GARCH (1, 1) family model is the most convenient specification in the financial literature [3, 13] to fit data with parsimonious model. But in this study we considered higher order models. As a result, the GARCH (1,1) family model is compared to various higher-order models from 1 and 2 or two months relationship of volatilities, since GARCH family model was used for short-term forecasting based on the value of AIC and BIC. In our model selection procedure we first fit different GARCH family models of different orders of p and q. As we can see from the table GARCH(0,1) model under all distributional assumptions (under normal distribution, t-distribution and GED distribution), EGARCH(1, 1) model under normal distributional assumption for residuals were selected as candidate models for the domestic price volatility of Teff, since they possesses minimum AIC and/or BIC. Moreover, to select an appropriate conditional volatility model, we consider the forecasting performance of the selected GARCH family models. The forecasting performance of the fitted GARCH family models are evaluated by RMSE, MAE, MAPE and Theil inequality coefficients. The model with the smallest statistics is considered to be better fit for modeling the conditional volatility of the domestic price of Teff in Ethiopia. The results are summarized in Table 8.

3.7.2. Parameter Estimation

Table 7: ML parameter estimation of EGARCH (1, 1) model under student t-distribution

Variables		Coefficient	Standard error	z-statistics	p-values
Mean equations	Constant	17.9874	11.2829	4.5942	0.0147
	AR(1)	0.6974	0.08059	8.6536	0.000
	MA(1)	-1.1592	0.0704	-16.4571	0.000
Variance equations					
		Coefficient	Standard error	z-statistics	p-values
	Constant	1.3180	0.1668	7.9004	0.000
	ARCH(-1)	-2.8249	1.0162	-2.7797	0.0054
	GARCH(-1)	-2.0585	1.1902	-1.7294	0.0037
	Asymmetric (-1)	0.9336	0.0086	107.9071	0.000
	Fuel oil	-0.3449	0.01466	-2.35	0.009
	Exchange rate	0.03651	0.00525	6.95	0.000
	Food	-0.0005	0.0001	-5.41	0.000
	Non-food	0.00276	0.0092	10.30	0.003
	Constant	2.8102	0.3574	7.86	0.000

EGARCH (1, 1) model with Normal distributional assumption for residuals were selected as better fit models based on forecasting accuracy measures, then the next step is estimation of parameters to perform analysis of the determinants of average monthly domestic price volatility of coffee. The parameters in the mean and variance equations are estimated using the maximum likelihood (ML) estimation method. The results are shown in Tables-7.

The results from the variance equation above show those among the explanatory variables which are considered in this study the coefficients of price of fuel oil and food price are negative and statistically significant at 1% level of significance. This shows that fuel oil price and food price has no significant impact on domestic price volatility of Teff. Alternatively the coefficients of price volatility of Teff, exchange rate (in birr against US dollar) and Non-food inflation rate are positive and statistically significant at 1% level of significance. That is price of Teff and non-food inflation rate has a positive significant effect on the current monthly domestic price volatility of Teff. This indicates that an increase in price volatility of Teff and non-food inflation rate leads to an increase in the monthly domestic price volatility of Teff. The results also show that the coefficient of 1 month lagged shocks (i.e ARCH (-1) GARCH (-1) terms) were negative and statistically significant at the 1% level. This shows that the current monthly domestic price volatility of Teff was affected by its 1-month lagged shocks. Additionally, the coefficient of the asymmetric term was positive and statistically significant at 1% level of significance. This indicating that bad news (unexpected increase in monthly domestic price of Teff) had larger impact on the domestic price volatility of Teff than good news (unexpected decrease in monthly domestic price volatility of Teff).

3.7.3. Model Adequacy Checking

In order to check whether the fitted model is a good fit to the data or not, that is the presence of remaining ARCH effect in the residuals, ARCH LM test and Ljung-Box Q-tests were performed. The Ljung-Box test indicates that the autocorrelations in the standardized residuals are not significantly different from zero for the first 24 lags (Tables-8). As can be seen from Table-8, the Breusch–Godfrey Serial Correlation LM test indicates that the standardized residuals of the fitted model did not exhibit any additional ARCH effect.

Table 8: ARCH-LM test for standardized residuals of the fitted volatility model Heteroskedasticity Test:
Breusch-Pagan-Godfrey

F-statistic	0.653074	Prob. F(1,85)	0.4213
Obs*R-squared	0.663344	Prob. Chi-Square(1)	0.4154

Source: author's estimation using Eviews-8

3.7.4. Forecasting price volatility

One of the fundamental uses of developing GARCH family models is forecasting. The volatility of domestic prices using variance as volatility measures was forecasted using the in-sample observations under statistic forecasting. The results are displayed as we can see from figure-4.

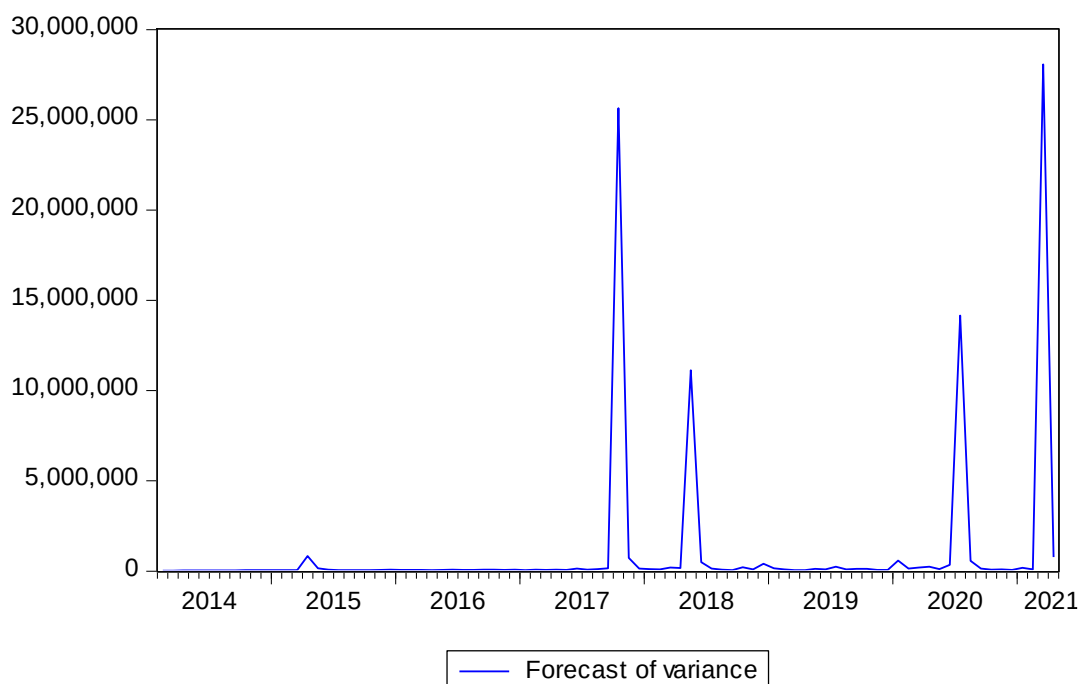


Figure 4: In-sample forecast of monthly domestic price volatility of Teff

Figure 4 In-sample forecast of monthly price volatility of Teff

As we can see from Figure 2, the domestic price volatility of Teff remained at almost a constant level at up to August, 2017 and increased steadily with increasing rate from the years end of August 2017 to January 2018, immediately drops sharply with increasing rate till March 2018 up to June 2018 and remained almost constant up to the end of the study. Moreover, high domestic price volatility was observed around the year February 2021.

4. Conclusion and recommendations

4.1. Conclusion

The study investigates forecasting average domestic price volatility of Teff in Ethiopia over the study periods from January 2014 to June 2021. Generalized Autoregressive Conditionally Heteroskedasticity (GARCH) family models including both symmetric models and asymmetric models were considered. The in-sample forecast using the best-fit EGARCH (1, 1) indicates that the price volatility of Teff remained at almost a constant level at up to August, 2017 and increased steadily with increasing rate from the year's end of August 2017 to January 2018. High price volatility was observed around the year February 2021.

4.2. Recommendation

The domestic price volatility of Teff was influenced by macroeconomic factors such as the exchange rate, food inflation rate, non-food inflation rate and fuel oil. Therefore, the government and concerned bodies should give due attention to these macroeconomic factors during policy formulation. Further studies should employ

multivariate models such as seasonal or dynamic conditional correlation multivariate models to analyze the time-varying correlation of the domestic price of Teff with other variables.

Reference

- [1]. Amentae TK, Tura EG, Gebresenbet G, Ljungberg D.,2016. Exploring value chain and post harvest losses of Teff in Bacho and Dawo districts of central Ethiopia. *J Stored Prod Postharvest Res* 2016; 7(1): 11-28.
- [2]. Bachewe FN, Koru B, Taffesse AS. 2015. Productivity and efficiency of smallholder teff farmers in Ethiopia International Food Policy Research Institute.
- [3]. Bollerslev, T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307-327.
- [4]. Bollerslev, T. 2009. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31:307-327
- [5]. Box, G.E.P. and Jenkins, G.M., 1976. *Time Series Analysis, Forecasting and Control*. Revised Edition, Holden Day.
- [6]. Dickey, D.A. and Fuller, W.A. 1979. Distributions of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* 74: 427-431.
- [7]. Efa Gobena, Degye Goshu, Tinsae Demisie and TadesseKenea. 2016. Determinants of market participation and intensity of marketed surplus of teff producers in Bacho and Dawo Districts of Oromia State, Ethiopia. *Journal of Agricultural Economics and Development*, 5(2): 20–32.
- [8]. Engle, R. 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrical*, **50**, 987–1007.
- [9]. Fantu Nisrane Bachewe, Bethelihem Koru, and Alemayehu Seyoum Taffesse.2015. Productivity and efficiency of smallholder teff farmers in Ethiopia, No.79 International Food Policy Research Institute (IFPRI).
- [10]. FAO., 2015. Analysis of price incentives for Teff in Ethiopia Technical notes series, MAFAP, by Assefa B. Demeke M., Lanos B, 2015 Rome.
- [11]. Fryzlewicz, P. 2007. Lecture Notes: Financial time series ARCH and GARCH models, Department of Mathematics, University of Bristol, Bristol BS8 1TW, UK. p.z.fryzlewicz@bristol.ac.uk, <http://www.maths.bris.ac.uk/~mapzf/>.
- [12]. Jordaan, H., Jooste, B.A. and Alemu, Z.G. 2007. Measuring the Price Volatility of Certain Field Crops in South Africa: Using the ARCH/GARCH Approach. *Journal of agriculture* 46.
- [13]. Lee, S.W., Hansen, B.E., 1994. Asymptotic properties of the maximum likelihood estimator and test of the stability of parameters of the GARCH and IGARCH models. *Econometric Theory* 10: 29-52.
- [14]. Marno Verbeek. 2004. *A Guide to Modern Econometrics* second Edition. Erasmus University, Rotterdam: John Wiley and Sons, Ltd.
- [15]. Mebrahatom Medhane, 2014. Determinants of commercialization of teff and its factor productivity outcome: the case of TahtayQoraroworeda, northwest zone of Tigray regional state, Ethiopia. MSc Thesis, Haramaya University Haramaya, Ethiopia.
- [16]. MelkamAnteneh. 2015. Teff (*Eragrostis tef* (zucc.) seed quality variation in East GojamZone, Ethiopia.

Malaysian Journal of Medical and Biological Research,2: 54-62

- [17]. Minten B, Tamru S, Engida E, Kuma T. 2016. Feeding Africa's cities: The case of the supply chain of teff to Addis Ababa. *Econ Dev Cult Change* 2016; 64(2): 265-97. [<http://dx.doi.org/10.1086/683843>]
- [18]. Phillips, C.B. and Perron, P.,1987. Testing for a Unit Root in Time Series Regression. *Biometrics* 75: 335-346.
- [19]. Shewartz, G. W. 1989. Why Does Stock Market Volatility Change over Time? *Journal of Finance* 44(5): 1115-1153.
- [20]. Tsay, R.S., 2005. *Analysis of Financial Time Series*, 2nd Edition. John Wiley and Sons, New York.
- [21]. Tsay, R.S. 2010. *Analysis of Financial Time Series*, Third Edition. New Jersey, John Wiley & Sons, Inc.