INTERNATIONAL JOURNAL OF MATHEMATICAL ANALYSIS AND
OPTIMIZATION: THEORY AND APPLICATIONS
VOL. 2018, PP. 291 - 311

On Optimal Estimate Functions for Asymmetric GARCH Models.

Onyeka-Ubaka, J. N[1*], Abass, O.[2]

1. *Department of Mathematics, University of Lagos, Nigeria*
2. *Department of Mathematics and Computer, Bells University, Ota, Nigeria*

** Corresponding author: e-mail: jonyeka2017@gmail.com*

Article Info
Received: 1 May 2018 Revised: 30 May 2018
Accepted: 16 July 2018 Available online: 26 July 2018

Abstract

High frequency data exhibit non-constant variance. This paper models the exhibited fluctuations via asymmetric GARCH models. The Maximum Likelihood Estimation (MLE) and Estimating Functions (EF) are used in the estimation of the asymmetric GARCH family models. This EF approach utilizes the third and fourth moments which are common features in financial time series data analysis and does not rely on distributional assumptions of the data. Optimal estimating functions have been constructed as a combination of linear and quadratic estimating functions. The results show that estimates from the estimating functions approach are better than those of the traditional estimation methods such as the MLE especially in cases where distributional assumptions on the data are seriously violated. The implementation of the EF approach to asymmetric GARCH models assuming a generalized student-t distribution innovation reveals the efficiency benefits of the EF approach over the MLE method in parameter estimation especially for non-normal cases.

Keywords: *Volatility clustering, Impact curve, Heavy tails, Asymmetric GARCH, EF*

MSC 2010: 37M10; 62M10; 91B25; 91B79; 91B84

1 Introduction

Financial market volatility can have a wide repercussion on the economy as a whole. The incidents caused by the terrorists attack of Boko Haram in the North East of Nigeria, some nefarious attitude of herdsmen on farmers in some parts of the country and the recent financial scandals of some political office holders in Nigeria have caused great turmoil in financial markets on several countries and a negative impact on the world economy ([1]Achumba, Ighomereho and Akpor-Robaro, 2013; [2]Mukolu and Ogodor, 2018). This is clear evidence of the important link between financial market uncertainty and public confidence. For this reason, policy makers often rely on market estimates of volatility as a barometer for the vulnerability of financial markets and the economy. High frequency financial time series data exhibit certain patterns which are crucial for correct model specification, estimation and forecasting. The stylized facts are fat tails, volatility clustering, leverage effects, long memory and co-movements in volatility ([3]Onyeka-Ubaka, 2013). The GARCH family of models has proved to be successful in capturing volatility clustering and some amount of the excess kurtosis which characterize financial time series data. Asset prices are generally non stationary. Some financial time series are fractionally integrated. Return series usually show no or little autocorrelation. Serial independence between the squared values of the series is often rejected pointing towards the existence of non-linear relationships between subsequent observations ([3]Onyeka-Ubaka, 2013; [4]Onyeka-Ubaka, Abass and Okafor, 2014; [5] Ogbogbo, 2018). Volatility of the return series appears to be clustered and normality has to be rejected in favour of some thick-tailed distribution ([6]Storti and Vitale, 2003). Some series exhibit so-called leverage effect, that is, changes in stock prices tend to be negatively correlated with changes in volatility. A firm with debt and equity outstanding typically becomes more highly leveraged when the value of the firm falls. This raises equity returns volatility if returns are constant. [7]Black, Fischer and Myron (1973) however, argued that the response of stock volatility to the direction of returns is too large to be explained by leverage alone. Volatilities of different securities very often move together. The works of [8]Engle (1982), [9]Bollerslev (1986) and various variants of the GARCH model have been developed to model volatility. Of great importance is the asymmetric GARCH family of models which address a major limitation of the [9]Bollerslev's (1986) basic GARCH model, relating to the inability of this model to capture the asymmetric impact of news on volatility. That is, there exists a negative correlation between stock returns and volatility implying that negative returns tend to be followed by larger increases in volatility while positive returns of the same magnitude tend to be followed by lower volatility. Different volatility models that capture this aspect have been proposed and widely applied to real life problems in the last two decades. Some of the most popular models include the EGARCH [10]Nelson (1991), GJR-GARCH [11]Glosten, Jagannathan and Runkle (1993), NAGARCH [12]Engle and Ng (1993), APARCH [13]Ding, Granger and Engle (1993), TGARCH [14]Zakoian (1994) and the QGARCH [15]Sentana, 1995). The motivation and objective of this paper is to explore the different impact of positive and negative shocks of equal magnitude on stock market volatility using asymmetric models.

ARCH process ([16]Bollerslev, Engle and Nelson, 1994)

The process $\{\varepsilon_t(\theta_0)\}$ follows an Autoregressive Conditional Heteroskedasticity (ARCH) model if

$$\mathrm{E}_{t-1}[\varepsilon_t(\theta_0)] = 0 \qquad t = 1,\ 2,\ \cdots \tag{1.1}$$

and the conditional variance

$$\sigma_t^2(\theta_0) = \mathrm{var}_{t-1}[\varepsilon_t(\theta_0)] = \mathrm{E}_{t-1}[\varepsilon_t^2(\theta_0)] = 0 \qquad t = 1,\ 2,\ \cdots \tag{1.2}$$

depends non trivially on the σ -field generated by the past observations: $\{\varepsilon_{t-1}(\theta_0),\ \varepsilon_{t-2}(\theta_0),\ \cdots\}$.

GARCH (p, q) Model ([9]Bollerslev, 1986)

In order to model in a parsimonious way the conditional heteroskedasticity, [9]Bollerslev (1986) proposed the generalized ARCH model, i.e GARCH(p, q):

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2 \tag{1.3}$$

where $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \cdots + \alpha_q L^q$, $\beta(L) = \beta_1 L + \beta_2 L^2 + \cdots + \beta_p L^p$. *The* GARCH(1,1) is the most popular model in the empirical literature:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{1.4}$$

To ensure that the conditional variance is well defined in a GARCH (p, q) model, all the coefficients in the corresponding linear ARCH(∞) should be positive [17]Rossi (2004):

$$\sigma_t^2 = \left(1 - \sum_{i=1}^{p} \beta_i L_i\right)^{-1} \left[\omega + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2\right]$$

$$= \omega^* + \sum_{k=0}^{\infty} \phi_k \varepsilon_{t-k-1} \tag{1.5}$$

$\sigma_t^2 \geq 0$ if ω∗ ≥ 0 and all $\phi_k \geq 0$. The non-negativity of ω∗ and ϕ_k is also a necessary condition for the non negativity of σ_t^2 . In order to make ω∗ and $\{\phi_k\}_{k=0}^{\infty}$ well defined, assume that: (i) the roots of the polynomial $\beta(x) = 1$ lie outside the unit circle, and that ω ≥ 0, this is a condition for ω∗ to be finite and positive. (ii) $\alpha(x)$ and $1 - \beta(x)$ have no common roots. These conditions are establishing nor that $\sigma_t^2 \leq \infty$ neither that $\{\sigma_t^2\}_{t=-\infty}^{\infty}$ is strictly stationary. For the simple GARCH(1,1) almost sure positivity of σ_t^2 requires, with the conditions (i) and (ii), that [18]Nelson and Cao (1992), $\omega \geq 0,\ \alpha_i \geq 0,\ \beta_i \geq 0$. For the GARCH(1, q) and GARCH(2, q) models these constraints can be relaxed, e.g. in the GARCH(1, 2) model the necessary and sufficient conditions become: $\omega \geq 0,\ \alpha_1 \geq 0,\ 0 \leq \beta_1 < 1,\ \beta_1\alpha_1 + \alpha_2 \geq 0$.

For the GARCH(2, 1) model the conditions are: $\omega \geq 0,\ \alpha_1 \geq 0,\ \beta_1 \geq 0,\ \beta_1 + \beta_2 < 1,\ \beta_1^2 + \beta_2 \geq 0$.

These constraints are less stringent than those proposed by [9]Bollerslev (1986):

$$\omega \geq 0;\ \beta_i \geq 0\ i = 1,\ \cdots,\ p;\ \alpha_j \geq 0\ j = 1,\ \cdots,\ q$$

These results cannot be adopted in the multivariate case, where the requirement of positivity for $\{\sigma_t^2\}$ means the positive definiteness for the conditional variance-covariance matrix. The process $\{\varepsilon_t\}$ which follows a GARCH(p, q) model is a martingale difference sequence. In order to study second-order stationarity it is sufficient to consider that: $\mathrm{var}[\varepsilon_t] = \mathrm{var}[\mathrm{E}_{t-1}(\varepsilon_t)] + \mathrm{E}[\mathrm{var}_{t-1}(\varepsilon_t)] = \mathrm{E}[\sigma_t^2]$

and show that is asymptotically constant in time (it does not depend upon time). A process $\{\varepsilon_t\}$ which satisfies a GARCH(p, q) model with positive coefficient $\omega \geq 0,\ \alpha_i \geq 0$ $i = 1,\ \cdots,\ q;\ \beta_i \geq 0\ i = 1,\ \cdots,\ p$ is covariance stationary if and only if: $\alpha(1) + \beta(1) < 1$. This is a sufficient but non necessary condition for strict stationarity. Because ARCH processes are thick tailed, the conditions for covariance stationarity are often more stringent than the conditions for strict stationarity. [10]Nelson (1991) showed that when $\omega > 0$, $\sigma_t^2 < \infty$ almost surely and $\{\varepsilon_t, \sigma_t^2\}$ is strictly stationary if and only if $\mathrm{E}[\ln(\beta_1 + \alpha_1 z_t^2)] < 0$.

$$\mathrm{E}[\ln(\beta_1 + \alpha_1 z_t^2)] \le \ln[\mathrm{E}(\beta_1 + \alpha_1 z_t^2)] = \ln(\alpha_1 + \beta_1) \tag{1.6}$$

when $\alpha_1 + \beta_1 = 1$, the model is strictly stationary. $\mathrm{E}[\ln(\beta_1 + \alpha_1 z_t^2)] < 0$ is a weaker requirement than $\alpha_1 + \beta_1 = 1$.

Example: ARCH(1), with $\alpha_1 = 1,\ \beta_1 = 0,\ z_t \sim \mathrm{N}(0,1)$

$$\mathrm{E}[\ln(z_t^2)] \le \ln[\mathrm{E}(z_t^2)] = \ln(1)$$

It is strictly but not covariance stationary. The ARCH(q) is covariance stationary if and only if the sum of the positive parameters is less than one.

Forecasting with a GARCH(p, q) ([19]Engle and Bollerslev 1986):

$$\sigma_{t+k}^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t+k-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t+k-i}^2 \tag{1.7}$$

we can write the process in two parts, before and after time t:

$$\sigma_{t+k}^2 = \omega + \sum_{i=1}^{n} [\alpha_i \varepsilon_{t+k-i}^2 + \beta_i \sigma_{t+k-i}^2] + \sum_{i=k}^{m} [\alpha_i \varepsilon_{t+k-i}^2 + \beta_i \sigma_{t+k-i}^2] \tag{1.8}$$

where n = min{m, k-1} and by definition summation from 1 to 0 and from k > m to m both are equal to zero. Thus

$$\mathrm{E}_t[\sigma_{t+k}^2] = \omega + \sum_{i=1}^{n} [(\alpha_i + \beta_i)\mathrm{E}_t(\sigma_{t+k-i}^2)] + \sum_{i=k}^{m} [\alpha_i \varepsilon_{t+k-i}^2 + \beta_i \sigma_{t+k-i}^2] \tag{1.9}$$

In particular for a GARCH(1, 1) and k > 2:

$$\begin{aligned}
\mathrm{E}_t[\sigma_{t+k}^2] &= \omega + \sum_{i=0}^{k-2} (\alpha_i + \beta_i)^i \omega + (\alpha_1 + \beta_1)^{k-1} \sigma_{t+1}^2 \\
&= \omega \frac{[1 - (\alpha_1 + \beta_1)^{k-1}]}{[1 - (\alpha_1 + \beta_1)]} + (\alpha_1 + \beta_1)^{k-1} \sigma_{t+1}^2 \\
&= \sigma^2 [1 - (\alpha_1 + \beta_1)^{k-1}] + (\alpha_1 + \beta_1)^{k-1} \sigma_{t+1}^2 \\
&= \sigma^2 + (\alpha_1 + \beta_1)^{k-1} [\sigma_{t+1}^2 - \sigma^2]
\end{aligned}$$

When the process is covariance stationary, it follows that $\mathrm{E}_t[\sigma_{t+k}^2]$ converges to σ^2 as $k \to \infty$.

The GARCH(p, q) process characterized by the first two conditional moments:

$$\mathrm{E}_{t-1}[\varepsilon_t] = 0$$

$$\sigma_t^2 \equiv \mathrm{E}_{t-1}[\varepsilon_t^2] = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{1.10}$$

where ω ≥ 0, $\alpha_i \ge 0$ and $\beta_i \ge 0$ for all i and the polynomial $1 - \alpha(x) - \beta(x) = 0$ has d > 0 unit root(s) and max{p, q} − d root(s) outside the unit circle is said to be: integrated in variance of order d if ω = 0 and integrated in variance of order d with trend if ω > 0. The integrated GARCH(p, q) models, both with or without trend, are therefore part of a wider class of models with a property called persistent variance in which the current information remains important for the forecasts of the conditional variances for all horizon. So we have the integrated GARCH(p, q) model when (necessary condition) $\alpha(1)+\beta(1)=1$.

The EGARCH Model ([10]Nelson, 1991)

In the EGARCH(p, q) model (Exponential GARCH(p, q)) put forward by Nelson the σ_t^2 depends on both size and the sign of lagged residuals. This is the first example of asymmetric model:

$$\ln(\sigma_t^2)=\omega+\sum_{i=1}^{p}\beta_j\ln(\sigma_{t-i}^2)+\sum_{i=1}^{q}\alpha_i\left[\phi z_{t-i}+\psi\left(\left|z_{t-i}\right|-\mathrm{E}\left|z_{t-i}\right|\right)\right] \qquad (1.11)$$

The left hand side is the log of the variance series. This makes the leverage effect exponential and therefore the parameters ω, β_j, α_i are not restricted to be nonnegative, $\alpha_1=1$, $\mathrm{E}|z_t|=(2/\pi)^{\frac{1}{2}}$ when $z_t\sim NID(0,1)$. Let define $g(z_t)=\phi z_t+\psi[|z_t|-\mathrm{E}|z_t|]$ by construction $\{g(z_t)\}_{t=-\infty}^{\infty}$ is a zero-mean, i.i.d. random sequence. The components of $g(z_t)$ are ϕz_t and $\psi[|z_t|-\mathrm{E}|z_t|]$, each with mean zero. If the distribution of z_t is symmetric, the components are orthogonal, but not independent. Over the range $0<z_t<\infty$, $g(z_t)$ is linear in z_t with slope $\phi+\psi$, and over the range $-\infty<z_t<0$, $g(z_t)$ is linear with slope $\phi-\psi$. The term $\psi[|z_t|-\mathrm{E}|z_t|]$ represents a magnitude effect. If $\psi>0$ and $\phi=0$, the innovation in $\ln\sigma_{t+1}^2$ is positive (negative) when the magnitude of z_t is larger (smaller) than its expected value. If $\psi=0$ and $\phi<0$, the innovation in conditional variance is now positive (negative) when returns innovations are negative (positive). A negative shock to the returns which would increase the debt to equity ratio and therefore increase uncertainty of future returns could be accounted for when $\alpha_i>0$ and $\phi<0$.

[10]Nelson (1991) assumes that z_t has a generalized error distribution (GED) (exponential power family). The density of a GED random variable normalized is:

$$f(z;v)=\frac{v\exp[-(1/2)|z/\lambda|^v}{\lambda 2^{(1+1/v)}\Gamma(1/v)} \qquad -\infty<z<\infty,\ 0<v<\infty \qquad (1.12)$$

where $\Gamma(\cdot)$ is the gamma function, and

$$\lambda\equiv\left[\frac{2^{(-2/v)}\Gamma(\frac{1}{v})}{\Gamma(\frac{3}{v})}\right]^{\frac{1}{v}}$$

where v is a tail thickness parameter, z's distribution; υ = 2 (standard normal distribution); υ < 2 (thicker tails than the normal); υ = 1 (double exponential distribution); υ > 2 (thinner tails than the normal); υ = ∞ (uniformly distributed on $\left[-3^{\frac{1}{2}},\ 3^{\frac{1}{2}}\right]$ with this density, $\mathrm{E}|z_t| = \dfrac{\lambda 2^{\frac{1}{v}}\Gamma\left(\frac{2}{v}\right)}{\Gamma\left(\frac{1}{v}\right)}$.

The asymmetric Non linear ARCH (p, q) model ([19]Engle and Bollerslev, 1986):

$$\sigma_t^{\gamma} = \omega + \sum_{i=1}^{q} \alpha_i |\varepsilon_{t-i}|^{\gamma} + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^{\gamma} \tag{1.13}$$

$$\sigma_t^{\gamma} = \omega + \sum_{i=1}^{q} \alpha_i |\varepsilon_{t-i} - k|^{\gamma} + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^{\gamma}$$

for $k \neq 0$, the innovations in σ_t^{γ} will depend on the size as well as the sign of lagged residuals, thereby allowing for the leverage effect in stock return volatility.

[11]The Glosten - Jagannathan - Runkle model (1993):

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{i=1}^{q} (\alpha_i \varepsilon_{t-i}^2 + \gamma_i S_{t-i}^{-} \varepsilon_{t-i}^2) \tag{1.14}$$

where

$$S_t^{-} = \begin{cases} 1 & if\ \varepsilon_t < 0 \\ 0 & if\ \varepsilon_t \geq 0 \end{cases}$$

The asymmetric GARCH(p, q) model ([20]Engle, 1990):

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i (\varepsilon_{t-i} + \gamma)^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{1.15}$$

The QGARCH by [15]Sentana (1995)

The Quadratic GARCH model captures asymmetry and has the form:

$$\sigma_t^2 = \sigma^2 + \psi' x_{t-q} + x'_{t-q} \mathrm{A} x_{t-q} + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{1.16}$$

when $x_{t-q} = \left(\varepsilon_{t-1},\ \cdots,\ \varepsilon_{t-q}\right)'$. The linear term $(\psi' x_{t-q})$ allows for asymmetry. The off-diagonal elements of A accounts for interaction effects of lagged values of x_t on the conditional variance. The QGARCH nests several asymmetric models. The proliferation of GARCH models has inspired some authors to define families of GARCH models that would accommodate as many individual models as possible.

The Asymmetric Power ARCH ([13]Ding, Granger and Engle, 1993)

$$r_t = \mu + \varepsilon_t$$

$$\varepsilon_t = \sigma_t z_t \qquad z_t \sim \mathrm{N}(0,1) \tag{1.17}$$

$$\sigma_t^\delta = \omega + \sum_{i=1}^{q} \alpha_i \left(|r_{t-i}| - \gamma_i \varepsilon_{t-i} \right)^\delta + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^\delta$$

where $\omega > 0;\ \delta \ge 0;\ \alpha_i \ge 0\ i = 1,\cdots,q;\ -1 < \gamma_i < 1\ i = 1,\cdots,q;\ \beta_j \ge 0\ j = 1,\cdots,p$. This model imposes a Box-Cox transformation of the conditional standard deviation process and the asymmetric absolute residuals. The Box-Cox transformation for a positive random variable γ_t :

$$\gamma_t^{(\lambda)} = \begin{cases} \dfrac{\gamma_t^\lambda - 1}{\lambda} & \lambda \ne 0 \\ \log \gamma_t & \lambda = 0 \end{cases} \tag{1.18}$$

The asymmetric response of volatility to positive and negative "shocks" is the well known leverage effect. This generalized version of ARCH model includes seven other models as special cases.

(i) ARCH(q) model, just let δ = 2 and $\gamma_i = 0,\ i = 1,\cdots,q,\ \beta_j = 0,\ j = 1,\cdots,p$

(ii) GARCH(p, q) model just let δ = 2 and $\gamma_i = 0,\ i = 1,\cdots,q$

(iii) Taylor/Schwert's GARCH in standard deviation model just let δ = 1 and $\gamma_i = 0,\ i = 1,\cdots,q$

(iv) GJR model just let δ = 2.

The GARCH-in-mean (GARCH-M) proposed by [21]Engle, Lilien and Robins (1987) consists of the system:

$$y_t = \gamma_0 + \gamma_1 x_t + \gamma_2 g(\sigma_t^2) + \varepsilon_t \qquad \varepsilon_t \big| \Phi_{t-1} \sim \mathrm{N}(0, \sigma_\varepsilon^2) \tag{1.19}$$

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^{q} \alpha_j \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

where $y_t \equiv (r_t - r_f)$, where $(r_t - r_f)$ is the risk premium on holding the asset, then the GARCH-M represents a simple way to model the relation between risk premium and its conditional variance. This model characterizes the evolution of the mean and the variance of a time series simultaneously. The GARCH-M model therefore allows analyzing the possibility of time-varying risk premium. It turns out that:

$$y_t \big| \Phi_{t-1} \sim \mathrm{N}(\gamma_0 + \gamma_1 x_t + \gamma_2 g(\sigma_t^2), \sigma_t^2) \tag{1.20}$$

In applications, $g(\sigma_t^2) = \sqrt{\sigma_t^2}$, $g(\sigma_t^2) = \ln(\sigma_t^2)$ and $g(\sigma_t^2) = \sigma_t^2$ have been used. [6]Storti and Vitale (2003) proposed BL-GARCH model in Gaussian framework. [22]Diongue, Guegan and Wolff (2010) extended their works using elliptical noise to capture the leverage effect or negative correlation between asset returns and volatility. The BL-GARCH model is given as

$$y_t = \mu + e_t$$
$$\varepsilon_t^2 = \sigma_t^2 z_t^2$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{k=1}^{r} c_k \sigma_{t-k} \varepsilon_{t-k} \qquad (1.21)$$

where $\alpha_0 > 0;\ \alpha_i \geq 0\ i = 1, \cdots, q;\ \beta_j \geq 0\ j = 1, \cdots, p\ and\ c_k^2 < 4\alpha_i \beta_j$; p, q, r are non-negative integers with r = min(p, q), σ_t^2 is the conditional variance of the process $\{\varepsilon_t\}$ which only depends on past σ^2 and ε^2's; ε_t is a sequence of independent identically distributed elliptical random variables with mean zero and unit variance, D(0, 1); z_t is an independent, identically distributed random variable with mean zero and variance unity; v_t is the daily trading volume, which is used as a proxy variable for the current information flow to the market.

If $\underline{c} = \underline{0}$, the model (1.21) reduces to the state space representation of the GARCH model. In this sense, the bilinear generalized autoregressive conditional heteroskedasticity model is an asymmetric extension of the symmetric generalized autoregressive conditional heteroskedasticity model.

The News Impact Curve

News is a huge factor that affects stock prices and therefore measuring its impact on stock market volatility is a crucial area of research in financial theory. This news has asymmetric effects on volatility. In the asymmetric volatility models, good news and bad news have different predictability for future volatility. The news impact curve characterizes the impact of past return shocks on the return volatility which is implicit in a volatility model. Holding constant the information dated t-2 and earlier, we can examine the implied relation between ε_{t-1} and σ_t^2, with $\sigma_{t-i}^2 = \sigma^2\quad i = 1, \cdots, p$. This impact curve relates past return shocks (news) to current volatility. This curve measures how new information is incorporated into volatility estimates.

For the GARCH model the News Impact Curve (NIC) is centered on $\varepsilon_{t-1} = 0$.

GARCH(1,1): $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$

The news impact curve has the following expression: $\sigma_t^2 = \mathrm{A} + \alpha\varepsilon_{t-1}^2$ where $\mathrm{A} = \omega + \beta\sigma^2$.

In the case of EGARCH model the curve has its minimum at $\varepsilon_{t-1} = 0$ and is exponentially increasing in both directions but with different parameters.

EGARCH(1,1): $\ln\sigma_t^2 = \omega + \beta\ln\sigma_{t-1}^2 + \phi z_{t-1} + \psi\left(\left|z_{t-1}\right| - \mathrm{E}\left|z_{t-1}\right|\right)$

where $z_t = \varepsilon_t / \sigma_t$. The news impact curve is

$$\sigma_t^2 = \begin{cases} \mathrm{A}\exp\left[\frac{\phi+\psi}{\sigma}\varepsilon_{t-1}\right] & for\ \varepsilon_{t-1} > 0 \\ \mathrm{A}\exp\left[\frac{\phi-\psi}{\sigma}\varepsilon_{t-1}\right] & for\ \varepsilon_{t-1} < 0 \end{cases}$$

where $\phi < 0,\ \psi + \phi > 0,\ \mathrm{A} = \sigma^2\exp\left[\omega - \alpha\sqrt{\frac{2}{\pi}}\right]$.

The EGARCH allows good news and bad news to have different impact on volatility, while the standard GARCH does not. The EGARCH model allows big news to have a greater impact on volatility than GARCH model. EGARCH would have higher variances in both directions because the exponential curve eventually dominates the quadrature, [12]Engle and Ng (1993).

The asymmetric GARCH(1,1) ([20]Engle, 1990): $\sigma_t^2 = \omega + \alpha(\varepsilon_{t-1} + \gamma)^2 + \beta\sigma_{t-1}^2$

The NIC is $\sigma_t^2 = \mathrm{A} + \alpha(\varepsilon_{t-1} + \gamma)^2$ is asymmetric and centered at $\varepsilon_{t-1} = -\gamma$

where $\mathrm{A} \equiv \omega + \beta\sigma^2, \;\; \omega > 0, \;\; 0 < \beta < 1, \;\; \sigma > 0, \;\; 0 \le \alpha < 1.$

The Glosten-Jagannathan-Runkle model: $\sigma_t^2 = \omega + \alpha\varepsilon_t^2 + \beta\sigma_{t-1}^2 + \gamma S_{t-1}^{-1}\varepsilon_{t-1}^2$

$$S_{t-1}^{-} = \begin{cases} 1 & if\ \varepsilon_{t-1} < 0 \\ 0 & otherwise \end{cases}$$

The NIC is centered at $\varepsilon_{t-1} = -\gamma$

$$\sigma_t^2 = \begin{cases} \mathrm{A} + \alpha\varepsilon_{t-1}^2 & if\ \varepsilon_{t-1} > 0 \\ \mathrm{A} + (\alpha + \gamma)\varepsilon_{t-1}^2 & if\ \varepsilon_{t-1} < 0 \end{cases}$$

where $\mathrm{A} = \omega + \beta\sigma^2, \;\; \omega > 0, \;\; 0 \le \beta < 1, \;\; \sigma > 0, \;\; 0 \le \alpha < 1, \;\; \alpha + \beta < 1.$

2 Estimating Functions Approach

In the bulk of literature available for the family of GARCH models, the maximum likelihood estimation method has been the most preferred in parameter estimation due to its simplicity and desirable properties. However, this method is based on distributional assumptions which are often violated in practice and thus alternative parameter estimation approaches are required. An alternative method of estimation is based on the Estimating Functions (EF) approach of [23]Godambe (1985). This approach takes into account higher order moments while estimating unknown parameters and does not rely on any distributional assumptions on the data for optimality. The generalized student-t distribution with one skewness parameter and two-tailed parameters offers the study the potential to improve our ability to fit the data in the tail regions which are critical to the risk management and other financial economic application ([24]Onyeka-Ubaka, Abass and Okafor, 2016). This is because downward movement of the markets is followed by higher volatilities than upward movement of the same magnitude (see [25]Campbell and Mackinlay, 1997; [26]Pagan and Schwert, 1990; [27]Locke and Sayers, 1993; [28]Muller and Yohai, 2002; [29]Eraker, Johannes and Polson, 2003).

We adapt optimal estimating functions' application to asymmetric GARCH family of models of [30]Mutunga, Islam and Orawo (2014). [31]Godambe and Thompson (1989) extended the concept of optimality of [23]Godambe's (1985) EF into a general setting using a more flexible conditioning method which is related to the concept of quasi-likelihood approach. Taking Y as an arbitrary sample space, they considered the class of EFs M_i which is a real function defined on $\mathrm{Y} \times \Theta$ such that

$$\mathrm{E}[\mathrm{M}_i(y_1, \; y_2, \; \cdots, \; y_n) | \mathrm{Y}_i] = 0 \tag{2.1}$$

where Θ is the parameter space and $\Upsilon_i(1=1,2,\cdots,k)$ is the σ-field generated by a specified partition on the sample space, Υ.

2.1 Parameter Estimation Using the Estimating Functions Approach

To estimate parameters of the EGARCH and GJR-GARCH models in a regression model set up using the EFs approach, optimal estimating functions approach to discrete time stochastic processes by [31]Godambe and Thompson (1989) was applied. Consider a general expression of the EGARCH and GJR-GARCH models in a regression model set up without making any distributional assumptions for the errors,

$$\begin{aligned} y_t &= x_t\omega + \varepsilon_t \\ y_t \big| \psi_{t-1} &\sim (x_t\omega, \sigma_t^2) \end{aligned} \tag{2.2}$$

where ψ_{t-1} is the information set at time $t-1$, σ_t^2 follows either an EGARCH or GJR-GARCH process and the component x_t could be composed of exogenous variables and/or lagged variables of the variable y_t which is a discrete time series process. Consider the first EGARCH model:

$$\begin{aligned} \varepsilon_t &= z_t\sigma_t \\ \ln\sigma_t^2 &= \omega + \beta\ln\sigma_{t-1}^2 + \phi z_{t-1} + \psi\big[|z_{t-1}| - \mathrm{E}|z_{t-1}|\big] \\ \sigma_t^2 &= \exp\big\{\omega + \beta\ln\sigma_{t-1}^2 + \phi z_{t-1} + \psi\big[|z_{t-1}| - \mathrm{E}|z_{t-1}|\big]\big\} \end{aligned} \tag{2.3}$$

where $\alpha_1 = 1$, $g(z_t) = \phi z_t + \psi\big[|z_t| - \mathrm{E}|z_t|\big]$ and z_t is an independent and identically distributed sequence of random variables. Let $\theta_1 = (\omega,\beta,\phi,\psi)$. We seek to estimate the unknown parameter vectors ω and θ_1 in the regression model (2.2) where σ_t^2 is as defined in (2.3). Similarly consider the first order GJR-GARCH model:

$$\begin{aligned} \varepsilon_t &= z_t\sigma_t \\ \sigma_t^2 &= \omega + \alpha\varepsilon_{t-1}^2 + \gamma S_{t-1}^- \varepsilon_{t-1}^2 + \beta\ln\sigma_{t-1}^2 \\ S_t^- &= \begin{cases} 1 & \textit{when } \varepsilon_t < 0 \\ 0 & \textit{otherwise} \end{cases} \end{aligned} \tag{2.4}$$

where z_t is an independent and identically distributed sequence of random variables. Let $\theta_2 = (\omega,\alpha,\beta,\gamma)$. Similarly we seek to estimate the unknown parameter vectors ω and θ_2 in the regression model (2.2) where σ_t^2 is as defined in (2.4).

To evaluate the optimal estimates of ω and θ_i (i = 1, 2) in each case, [31]Godambe and Thompson's (1989) theorem for stochastic processes is applied. A good combination for basic unbiased and mutually orthogonal EFs is λ_{it} and λ_{2t} such that

$$\begin{aligned} \lambda_{it} &= y_t - x_t\omega \\ \lambda_{2t}^* &= (y_t - x_1\omega)^2 - \sigma_t^2 \end{aligned} \tag{2.5}$$

The choice of these two estimating functions is based on the need to estimate the conditional mean $x_t\omega$ and conditional variance σ_t^2 of y_t simultaneously. However the EF λ_{2t}^* is not orthogonal to the EF λ_{1t}. This implies that

$$E(\lambda_{1t}\lambda_{2t}^*) \neq 0 \tag{2.6}$$

$$\begin{aligned} E(\lambda_{1t}\lambda_{2t}^*) &= E\{(y_t - x_t\omega)[(y_t - x_t\omega)^2 - \sigma_t^t]\psi_{t-1}\} \\ &= E\{(y_t - x_t\omega)^3 - \sigma_t^2(y_t - x_t\omega)|\psi_{t-1}\} \\ &= E\{(y_t - x_t\omega)^3|\psi_{t-1}\} \neq 0 \end{aligned}$$

λ_{2t}^* is therefore orthogonalised using the Gram-Schmidt orthognalisation procedure as follows

$$\begin{aligned} \lambda_{2t} &= (y_t = x_t\omega)^2 - \sigma_t^2 \\ &= \{E[(y_t - x_1\omega)^2 - \sigma_t^2(y_t - x_t\omega)|\psi_{t-1}]E[(y_t - x_t\omega)^2|\psi_{t-1}]^{-1}(y_t - x_t\omega)|\psi_{t-1}\} \\ &= (y_t - x_t\omega)^2 - \sigma_t^2 - (y_t - x_t\omega)E\left[\frac{(y_t - x_t\omega)}{\sigma_t^2}\right]^2|\psi_{t-1} \end{aligned} \tag{2.7}$$

From (2.7), consider the component

$$E\left\{\frac{(y_t - x_t\omega)^2}{\sigma_t^2}|\psi_{t-1}\right\} \tag{2.8}$$

Dividing and multiplying (2.8) by σ_t we have

$$E\left\{\frac{(y_t - x_t\omega)^2}{\sigma_t^{3/2}}|\psi_{t-1}\right\}\sigma_t = \phi_{1t}\sigma_t \tag{2.9}$$

where ϕ_{1t} is the skewness of y_t conditional on ψ_{t-1}. Thus

$$\lambda_{2t} = (y_t - x_t\omega)^2 - \sigma_t^2 + \varepsilon_t\phi_{1t}\sigma_t \tag{2.10}$$

Therefore the two elementary EFs are now orthogonal as:

$$\begin{aligned} \lambda_{1t} &= y_t - x_t\omega \\ \lambda_{2t} &= (y_t - x_t\omega)^2 - \sigma_t^2 + \varepsilon_t\phi_{1t}\sigma_t \end{aligned} \tag{2.11}$$

To estimate the coefficient vectors ω and θ in the regression model (2.2), optimal EFs are derived using the elementary EFs in (2.11). The theorem by [31]Godambe and Thompson (1989) is applied to form a linear combination of the elementary EFs as

$$\begin{aligned} g_1 &= \sum_{t=1}^{T} a_{1t}\lambda_{1t} + \sum_{t=1}^{T} a_{2t}\lambda_{2t} \\ g_2 &= \sum_{t=1}^{T} b_{1t}\lambda_{1t} + \sum_{t=1}^{T} b_{2t}\lambda_{2t} \end{aligned} \tag{2.12}$$

Let L be the class of all EFs (g_1, g_2) given by (2.12). The jointly optimal EFs (g_1^*, g_2^*) are given by (2.8) with $a_{1t} = a_{1t}^*$ and $b_{1t} = b_{1t}^*$ for $i = 1, 2$ $t = 1, 2, \ldots$ T, where,

$$b_{1t}^* = \frac{\mathrm{E}\left(\frac{\partial \lambda_{1t}}{\partial \omega}|\psi_{t-1}\right)}{\mathrm{E}\left(\lambda_{1t}^2|\psi_{t-1}\right)} = \frac{\frac{\partial x_t \omega}{\partial \omega}}{\sigma_t^2} \tag{2.13}$$

$$b_{2t}^* = \frac{\mathrm{E}\left(\frac{\partial \lambda_{2t}}{\partial \omega}|\psi_{t-1}\right)}{\mathrm{E}\left(\lambda_{2t}^*|\psi_{t-1}\right)} = \frac{\mathrm{E}\frac{\partial}{\partial \omega}\left[(y_t - x_t\omega)^2 - \sigma_t^2 - (y_t - x_t\omega)\phi_1\sigma_t|\psi_{t-1}\right]}{\mathrm{E}\left(\lambda_{2t}^2|\psi_{t-1}\right)} \tag{2.14}$$

Solving the numerator in (2.14) we have

$$\mathrm{E}\frac{\partial}{\partial \omega}\left[(y_t - x_t\omega)^2 - \sigma_t^2 - (y_t - x_t\omega)\phi_1\sigma_t|\psi_{t-1}\right]$$

$$= \mathrm{E}\left\{-2(y_t - x_1\omega)\frac{\partial x_t\omega}{\partial \omega} - \frac{\partial \sigma_t^2}{\partial \omega} - \frac{(y_t - x_t\omega)}{2\sigma_t}\frac{\partial \sigma_t}{\partial \omega}\phi_{1t} + \phi_1\sigma_t\frac{\partial x_t\omega}{\partial \omega}\right\} \tag{2.15}$$

$$= \phi_{1t}\sigma_t\frac{\partial x_t\omega}{\partial \omega} - \frac{\partial \sigma_t^2}{\partial \omega}$$

Solving the denominator in (2.15) we have

$$\mathrm{E}\left(\lambda_{2t}^2|\psi_{t-1}\right) = \mathrm{E}\left\{\left[\begin{array}{l}(y_t - x_t\omega)^4 - 2\sigma_t^2(y_t - x_t\omega)^2 + \sigma_t^4 + \phi_{1t}\sigma_t^{\frac{3}{2}}(y_t - x_t\omega) - \phi_{1t}\sigma_t(y_t - x_t\omega)^3 \\ + \phi_{1t}\sigma_t^{\frac{3}{2}}(y_t - x_t\omega) - \phi_{1t}\sigma_t(y_t - x_t\omega)^3 + \phi_{1t}^2\sigma_t^2(y_t - x_t\omega)^2\end{array}\right]|\psi_{t-1}\right\} \tag{2.16}$$

Multiplying and dividing (2.16) by σ_t^2 leads to

$$= \mathrm{E}\sigma_t^4\left\{\left[\begin{array}{l}\frac{(y_t - x_t\omega)^4}{\sigma_t^4} - \frac{2\sigma_t^2(y_t - x_t\omega)^2}{\sigma_t^4} + 1 + \frac{\phi_{1t}(y_t - x_t\omega)}{\sigma_t} - \frac{\phi_{1t}\sigma_t(y_t - x_t\omega)^3}{\sigma_t^{\frac{3}{2}}} \\ + \frac{\phi_{1t}(y_t - x_t\omega)}{\sigma_t} - \frac{\phi_{1t}\sigma_t(y_t - x_t\omega)^3}{\sigma_t^{\frac{3}{2}}} + \frac{\phi_{1t}^2(y_t - x_t\omega)^2}{\sigma_t^2}\end{array}\right]|\psi_{t-1}\right\} \tag{2.17}$$

$$= \sigma_t^4\left(\phi_{2t} + 2 - \phi_{1t}^2\right)$$

where $$\phi_{2t} = \mathrm{E}\left\{\left[\frac{(y_t - x_t\omega)^4}{\sigma_t^4} - 3\right]|\psi_{t-1}\right\} \tag{2.18}$$

Equation (2.18) represents the standardized conditional kurtosis (excess kurtosis). Hence,

$$b_{2t}^* = \frac{\phi_{1t}\sigma_t\frac{\partial x_t\omega}{\partial \omega} - \frac{\partial \sigma_t^2}{\partial \omega}}{\sigma_t^4\left(\phi_{2t} + 2 - \phi_{1t}^2\right)} \tag{2.19}$$

$$a_{1t}^* = \frac{\mathrm{E}\left(\frac{\partial \lambda_{1t}}{\partial \theta}|\psi_{t-1}\right)}{\mathrm{E}\left(\lambda_{1t}^2|\psi_{t-1}\right)} = 0 \tag{2.20}$$

$$a_{2t}^* = \frac{\mathrm{E}\left(\frac{\partial \lambda_{2t}}{\partial \theta}|\psi_{t-1}\right)}{\mathrm{E}\left(\lambda_{2t}^2|\psi_{t-1}\right)}$$

$$= \frac{\mathrm{E}\frac{\partial}{\partial\theta}\left[(y_t - x_t\omega)^2 - \sigma_t^2 - (y_t - x_t\omega)\phi_{1t}\sigma_t\right]\psi_{t-1}}{\mathrm{E}\left(\lambda_{2t}^2 \middle| \psi_{t-1}\right)}$$

$$= \frac{\mathrm{E}\left[\frac{\partial\sigma_t^2}{\partial\theta} - \frac{1}{2}\phi_{1t}\frac{(y_t - x_t\omega)}{\sigma_t}\frac{\partial\sigma_t^2}{\partial\theta}\right]\Big|\psi_{t-1}}{\mathrm{E}\left(\lambda_{2t}^2 \middle| \psi_{t-1}\right)}$$

$$= \frac{-\frac{\partial\sigma_t^2}{\partial\theta}}{\sigma_t^4\left(\phi_{2t} + 2 - \phi_{1t}\right)} \tag{2.21}$$

Substituting (2.13), (2.19), (2.20) and (2.21) into (2.12) gives the jointly optimal EFs as

$$g_1^* = -\sum_{t=1}^{\mathrm{T}} \frac{\frac{\partial\sigma_t^2}{\partial\theta}}{\sigma_t^4\left(\phi_{2t} + 2 - \phi_{1t}\right)}\lambda_{2t}$$
$$g_2^* = \sum_{t=1}^{\mathrm{T}} \frac{\frac{\partial x_t\omega}{\partial\omega}}{\sigma_t^2}\lambda_{1t} - \sum_{t=1}^{\mathrm{T}} \frac{\frac{\partial\sigma_t^2}{\partial\theta}}{\sigma_t^4\left(\phi_{2t} + 2 - \phi_{1t}\right)}\lambda_{2t} \tag{2.22}$$

where σ_t^2 is given by equations (2.3) and (2.4) and

$$\theta = \begin{cases} \theta_1 & for\ EGARCH(1,1) \\ \theta_2 & for\ GJR - GARCH(1,1) \end{cases} \tag{2.23}$$

The result in (2.23) is very general in that no distributional assumptions on $y_t|\psi_{t-1}$ have been made. The estimates for the unknown parameter vectors ω and θ are obtained by solving the optimal EF in (2.23). This means numerically minimizing $g_1^* + g_2^*$.

$$g_{\theta,\omega}^* = g_1^* + g_2^* = 0 \tag{2.24}$$

where g_1^* and g_2^* are as defined in (2.22).

3 Results and Discussion

3.1 Data for the Study

The First Bank of Nigeria (FBN), Guaranty Trust Bank (GTB), United Bank for Africa (UBA) and Zenith Bank (ZEB) data sourced from Central Bank of Nigeria Annual Bulletin are selected for study from 4th January, 2007 to 25th November, 2017. The period is so selected because it followed the period, that the world economy grew steadily from 2002 to 2006 on the back of emerging economies of scale and interest rates remain low worldwide [32]United Nations (2006). Thanks to globalization that provides very accommodative monetary conditions. This nurtured expectation for stable growth and firm assets' prices to continue, result in a massive fund flow to money markets and a huge distortion in the market pricing of assets in October, 2008; and the Central Bank of Nigeria (CBN) intervention in consolidating and recapitalizing the banking sector through such process like the Expanded Discount Window (EDW). By means of purposive sampling technique, four banks were selected for the study. These four banks were selected because they are considered to be more susceptible to volatility than other banks and had passed the screening exercise conducted by CBN in August, 2009 ([3]Onyeka-Ubaka, 2013). Their volume of stocks traded on the floor of the Nigeria Stock Exchange (NSE) for the sample period were collected and analyzed using asymmetric GARCH models under the MLE and EF procedures. In each case daily returns are computed as logarithmic price (P_t) relatives.

$$r_t = \log \frac{\mathrm{P}_t}{\mathrm{P}_{t-1}} \tag{3.1}$$

where r_t is the log return series (continuously compounded return). Since $\hat{\mathrm{P}}_1$ depends on P_0 which is unknown, we ignore this quantity and begin the calculation of the fitted values with $\hat{\mathrm{P}}_2$.

3.2 Preliminary Tests

Table 1 presents summary statistics and preliminary tests of normality and asymmetry for the daily stock returns of the four financial series. We noticed that the daily volatility for the FBN index, represented by the standard deviation (2.74%) is above the volatility (1.95%, 1.69% and 1.83%) for the GTB, UBA and ZEB indices return series respectively.

Series Statistics	FBN	GTB	UBA	ZEB
Mean	-0.00059	-0.000908	-0.000061	-0.000792
Std Dev	0.027351	0.019540	0.016851	0.018346
Skewness	-0.173290	-2.385904	-1.031667	-0.168971
Kurtosis	5.743612	8.137659	7.045382	9.452281
Jarque-Bera (Probability)	0.0000	0.0000	0.0000	0.0000
σ	1	1	1	1
ADF test (Probability)	1.200E-3	1.000E-3	1.011E-3	1.0456E-3
σ	1	1	1	1
$\mathrm{Cov}(r_t^2, r_{t-1})$	-0.071853	-0.089452	-0.068723	-0.079530

Table 1: Summary Statistics of the compounded returns r_t

The skewness coefficients are negative for all series suggesting that they have long left tail. The kurtosis coefficients on the other hand are very high, a reflection that the distributions of the four sets of real data are highly leptokurtic. The p-value corresponding to the Jarque-Bera normality test is zero at 5% level suggesting that the test is significant for all series. The test gave a value of $\sigma = 1$ which indicates that the series r_t does not come from a normal distribution; in favour of $\sigma = 0$ which indicates that the series r_t comes from a normal distribution with unknown mean and variance. The test results imply that the four series exhibit non-normal behaviour. The Augmented Dickey-Fuller (ADF) test rejects the unit root null hypothesis in all data sets. This is indicated by the minimal p-values at 5% level and the values of σ. The test returns a value of $\sigma = 1$ which indicates rejection of the unit root in favour of the trend-stationary alternative. $\sigma = 0$ indicates failure to reject the unit root null. The series statistics show strong serial correlations in both levels of the return series. The results obtained are consistent with the results of [33]Pitt and Shephard (1999), [34]Storvik (2002), [35]Johannes, Polson and Stroud (2006), [36]Raggi and Bordignon (2006), who found that serial correlations in DJIA returns are significant but unstable and depend on the sample period. Thus we conclude that the returns of all stock indices are stationary. Finally, we performed the test for presence of asymmetric effects on conditional volatility in all empirical series. A simple diagnostic test for the leverage effects involves computing the sample correlation between squared returns and the lagged returns, Cov (r_t^2, r_{t-1}) ([37]Zivot, 2008). A negative value for this coefficient provides evidence for potential asymmetric effects. All series have a negative value for this coefficient indicating evidence of asymmetry and hence asymmetric GARCH family of models could perform well in explaining conditional volatility in this case.

Figure 1 presents the plot of daily logarithmic returns for the series over the sampled period. We observe that volatility clustering is present in the four cases as the four series show periods of low volatility which tend to be followed by periods of relatively low volatility and other periods of high volatility which likewise tend to be followed by high volatility as observed by [38]Mandelbrot (1963). This aspect can be thought of as clustering of the variance of the error term over time, that is, the error term exhibits time varying heteroskedasticity.

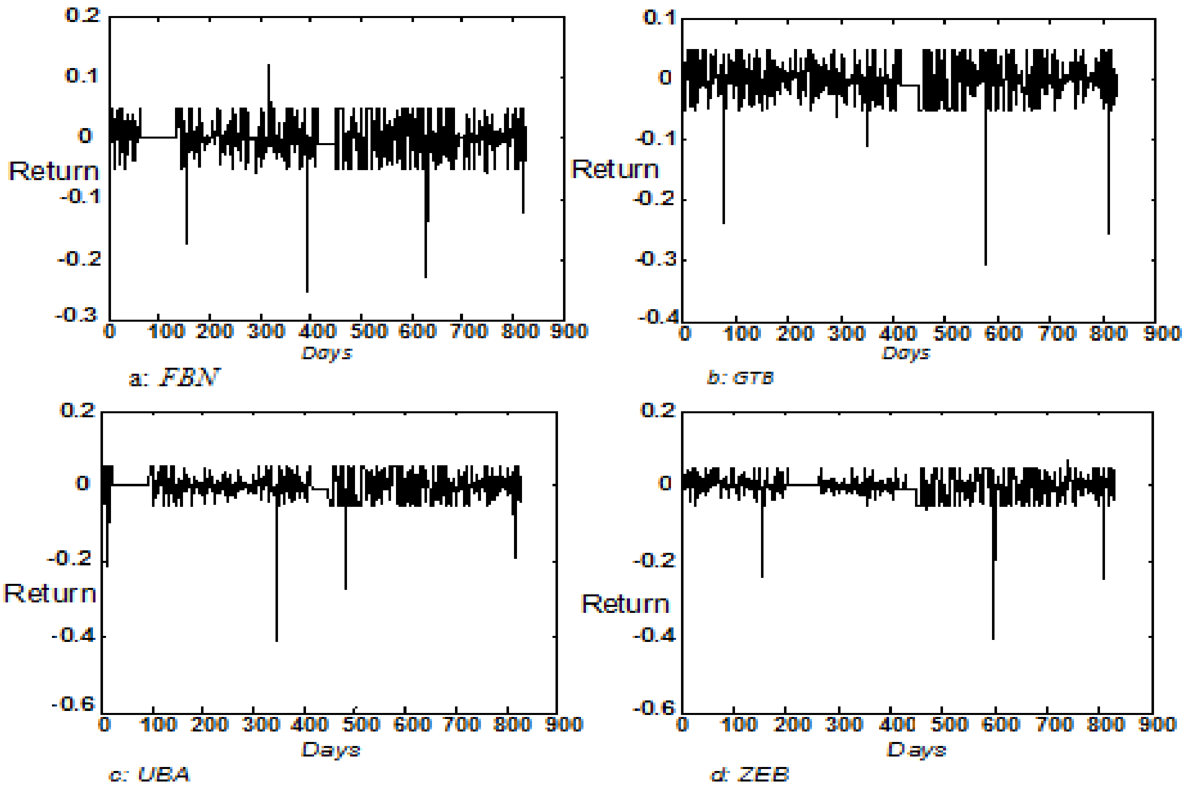

Fig. 1: Compounded Daily Returns of the Selected Four Banks (Jan., 2007 – Nov., 2017)

3.3 Model Estimation Results

The first order EGARCH and GJR-GARCH models are fitted to the four empirical series and estimates obtained using the maximum likelihood estimation and estimating function approaches. In parameter estimation under maximum likelihood method, we assume a standardized Gaussian or Generalized student-t distribution with $v = 10$ degrees of freedom for the innovations. Parameter estimates, corresponding standard errors (in parentheses), Akaike Information Criteria (AIC) and the log likelihood values are presented in Tables 2-4.

Series/Estimate Method		ω	β	ϕ	ψ
	MLE*	-0.127691	0.972876	0.173354	-0.262197
		(0.052301)	(0.007974)	(0.086452)	(0.070125)
FBN	MLE**	-0.214366	0.958671	0.149705	-0.178261
		(0.046812)	(0.006382)	(0.033415)	(0.027312)
	EF	-0.203167	0.958244	0.1492856	-0.175384
		(0.045352)	(0.004867)	(0.032918)	(0.026843)
	MLE*	-0.172631	0.977015	0.189403	-0.280332
		(0.103058)	(0.019546)	(0.076860)	(0.069115)
GTB	MLE**	-0.290344	0.965567	0.176915	-0.169378
		(0.068125)	(0.007138)	(0.038212)	(0.029436)
	EF	-0.293143	0.968134	0.169326	-0.171457
		(0.071532)	(0.007062)	(0.0367313)	(0.028810)
	MLE*	-0.150543	0.984593	0.170069	-0.273548
		(0.047332)	(0.006774)	(0.076635)	(0.083012)
UBA	MLE**	-0.208421	0.970115	0.157910	-0.183462
		(0.056172)	(0.005843)	(0.035782)	(0.0276791)
	EF	-0.203167	0.975633	0.162058	-0.179753
		(0.054850)	(0.005621)	(0.034924)	(0.0272814)
	MLE*	-0.190345	0.923487	0.169572	-0.269037
		(0.050923)	(0.008146)	(0.001535)	(0.076770)
ZEB	MLE**	-0.247889	0.940178	0.158731	-0.194782
		(0.074138)	(0.007195)	(0.036742)	(0.026953)
	EF	-0.253716	0.943812	0.155849	-0.189233
		(0.069105)	(0.007209)	(0.033759)	(0.026687)

*Standardized Gaussian distribution **Generalized Student's- t distribution (with degree of freedom $v = 10$)
Table 2: Parameter Estimates of EGARCH(1, 1) Model

From our parameter estimates it is clear that the EF method is more efficient than the MLE method in parameter estimation of the first order EGARCH and GJR–GARCH models in finite samples. The assertion is evident from the bracketed standard errors. The standard errors of the EF method estimates are smaller than those of the maximum likelihood estimates assuming either a Gaussian or a Generalized Student–t distribution with 10 degrees of freedom. The gain in efficiency follows from the fact that the EF method does not rely on distributional specifications for optimality and that it accounts for higher order moments present in non–normal data such as the present study financial time series. However, it is evident that the MLE method when assuming a Generalized Student-t error distribution competes reasonably well with the EF method and provides a better in-sample-fit than the MLE method when assuming a Gaussian error distribution across all data sets. This finding is expected considering the Jarque-Bera normality test in Table 1 which implies that the empirical distributions of the four return series exhibit heavier tails than the standard normal distribution. Moreover, a Generalized Student–t distribution exhibits excess kurtosis and fat tail behaviour.

Series/Estimate Method		ω	α	β	γ
FBN	MLE*	1.87412E-06 (1.30583E-06)	0.003196 (0.059174)	0.897354 (0.048716)	0.195734 (0.067135)
	MLE**	2.43951E-06 (1.23046E-06)	0.001264 (0.021637)	0.915479 (0.020354)	0.182961 (0.038271)
	EF	3.15926E-06 (1.197452E-06)	0.001185 (0.018943)	0.913285 (0.019518)	0.178354 (0.035618)
GTB	MLE*	1.28574E-05 (3.10345E-06)	0.001282 (0.062546)	0.886479 (0.050187)	0.217138 (0.070136)
	MLE**	1.19312E-05 (4.04186E-06)	0.003679 (0.027682)	0.946577 (0.019812)	0.189761 (0.043872)
	EF	1.221694E-05 (4.123794E-06)	0.006281 (0.025682)	0.943672 (0.016823)	0.185591 (0.037526)
UBA	MLE*	2.80367E-05 (1.92613E-06)	0.0021989 (0.051068)	0.860627 (0.035148)	0.207573 (0.054615)
	MLE**	3.85395E-05 (2.10798E-06)	0.0034976 (0.004293)	0.858491 (0.028354)	0.173296 (0.036729)
	EF	3.70159E-05 (2.09681E-06)	0.0031852 (0.003950)	0.854173 (0.027159)	0.168395 (0.033768)
ZEB	MLE*	1.92458E-06 (1.28673E-06)	0.0021348 (0.028714)	0.877215 (0.048716)	0.220547 (0.049670)
	MLE**	1.67213E-06 (1.18695E-06)	0.0041975 (0.037619)	0.903147 (0.019683)	0.212456 (0.036127)
	EF	1.59764E-06 (1.179148E-06)	0.0040871 (0.034712)	0.901582 (0.017069)	0.198191 (0.034918)

*Standardized Gaussian distribution **Generalized Student's- t distribution (with degree of freedom $v = 10$)

Table 3: Parameter Estimates of GJR-EGARCH (1, 1) Model

	EGARCH (1, 1)				***GJR-GARCH (1, 1)***			
Model Diagnostics	FBN	GTB	UBA	ZEB	FBN	GTB	UBA	ZEB
AIC MLE	2070.86	2308.51	3546.87	4618.19	2146.58	2316.59	3571.12	4625.03
EF	2071.17	2310.85	3548.36	4620.45	2147.11	2318.03	3572.23	4627.15
Log-Likelihood MLE	1839.47	1841.76	1832.69	1822.94	1839.21	1841.26	1832.08	1822.12
EF	1842.09	1843.37	1836.34	1827.26	1842.03	1843.52	1835.67	1827.07

Table 4: Model Evaluation Criteria

From the results, it is observed that both models have almost similar AIC and Log-likelihood values for the four financial series data. However, EGARCH (1,1) has relatively higher Log-likelihood and lower AIC values than GJR – GARCH (1,1) indicating that it performed relatively better in explaining conditional volatility in all empirical series over the considered study period. The coefficients σ^2 and σ for the first order EGARCH and GJR-GARCH models respectively reflects the leverage effects. The estimates indicate the magnitude and sign of the leverage effects. The EGARCH model shows a negative parameter of asymmetry in all financial series suggesting that past negative shocks (bad news) have a greater impact on subsequent volatility of returns than positive shocks (good news) do. The GJR-GARCH model records positive leverage effects, attesting that bad news in the market lead to a higher volatility of asset returns than good news.

4 Conclusion

In this paper, high frequency volatility data were modelled via asymmetric GARCH family models. The Maximum Likelihood Estimation and optimal Estimating Function methods were applied in parameter estimation. The results show that the EF method competes reasonably well with the MLE method especially in cases where there are serious departures from normality in finite samples. The paper extends and generalizes [30]Mutunga, Islam and Orawo (2014) by using generalized student-t distribution in asymmetric properties in high frequency volatility data. The present EF approach, therefore, provides a useful alternative method of estimation to the MLE method for the asymmetric GARCH models especially in cases where the true distribution of the data is unknown.

Competing financial interests

The authors declare no competing financial interests.

References

[1] Achumba, I. C., Ighomereho, O. S. & Akpor-Robaro, M. O. M. Security Challenges in Nigeria and the Implications for Business Activities and Sustainable Development. Journal of Economics and Sustainable Development, l4(2), 79-99 (2013).

[2] Mukolu, M. O & Ogodor, B. N. Insurgency and Implication on Nigeria economic growth. *International Journal of Development and Sustainability*, 7(2), 492-501 (2018).

[3] Onyeka-Ubaka, J. N. A BL-GARCH Model for Distribution with Heavy Tails. A Ph.D. *Thesis*, Department of Mathematics, University of Lagos, Nigeria (2013).

[4] Onyeka-Ubaka, J. N., Abass, O. & Okafor, R. O. Conditional Variance Parameters in Symmetric Models. *International Journal of Probability and Statistics*, 3(1), 1-7 (2014).

[5] Ogbogbo, C. P. Modeling crude oil spot price as an Ornstein- Uhlenbeck Process. *International Journal of Mathematical Analysis and Optimization: Theory and Applications*, 2018, 261-275 (2018).

[6] Storti, G. & Vitale, C. BL-GARCH models and asymmetries in volatility. *Statistical Methods and Applications*, 12, 19-40 (2003).

[7] Black, F. & Scholes, M. The Pricing of Options and Corporate Liabilities. *Journal of Political Economics,* 81, 637-54 (1973).

[8] Engle, R. F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987-1007 (1982).

[9] Bollerslev, T. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31, 307-328 (1986).

[10] Nelson, D.B. Conditional heteroskedasticity in Asset returns: A new Approach. *Econometrica*, 59, 347- 370 (1991).

[11] Glosten, L. R., Jagannathan, R. & Runkle, D. On the relation between the expected value and the volatility of nominal excess returns on stocks. *Journal of Finance*, 48: 1779-1801 (1993).

[12] Engle, R. F. & Ng, V. K. Measuring and testing the impact of news on volatility. *Journal of Finance*, 48(5), 1749-1778 (1993).

[13] Ding, Z., Granger, C. and Engle, R. A long memory property of stock returns and a new model. *Journal of Empirical Finance*, 1, 83–106 (1993).

[14] Zakoian, J.M. Threshold Heteroskedasticity Models. *Journal of Economic Dynamics and Control*, 19, 931-944 (1994).

[15] Sentana, E. Quadratic ARCH Models. *Review of Economic Studies*, 62, 639-661 (1995).

[16] Bolleslev, T., Engle, R. F. and Nelson, D. B. ARCH Models in Handbook of Econometrics, Vol. IV, R. F. Engle and D. L. McFadden, eds. Amsterdan: North Holland, 2959-3038 (1994).

[17] Rossi, E. Lecture Notes on GARCH Models. University of Pavia, 1, 33-45 (2004).

[18] Nelson, D. B. & Cao, C. Q. Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics*,10, 229-235 (1992).

[19] Engle, R. F. & Bollerslev T. Modelling the persistence of conditional variances, *Econometric Reviews*, 5, 1-50 (1986).

[20] Engle, R. F. Stock Volatility and the Crash of '87. *Review of Financial Studies*, 3(1), 103-106 (1990).

[21] Engle R. F., Lilien, D. M. & Robins, R. P. Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica,* 55(2), 391-407 (1987).

[22] Diongue, A. K., Guegan, D. & Wolff, R. C. Exact Maximum likelihood Estimator for the BL-GARCH model under elliptical distributed innovations. *Journal of Statistical Computation and Simulation*, 7, 4-10 (2010).

[23] Godambe, V.P. The foundations of finite sample estimation in stochastic processes, *Biometrika*, 72, 319-328 (1985).

[24] Onyeka-Ubaka, J. N., Abass, O. & Okafor, R. O. A Generalized t-Distribution based Filter for Stochastic Volatility Models, *AMSE JOURNALS-2016-Series:Advances D,* 21(1), 19-37 (2016).

[25] Campbell, J. Y. Lo, A. W. & Mackinlay, A. C. The Econometrics of Financial Markets. Princeton, New Jersey, Princeton University Press. *Journal of Finance*. 25, 383-417 (1997).

[26] Pagan, A. R. and Schwert, G. W. (1990). Alternative Models for Conditional Stock Volatility, *Journal of Economics*, 45, 267-290.

[27] Locke, P.R. & Sayers, C.L. Intra-Day Futures Prices Volatility: Information Effects and Variance Persistence. *Journal of Applied Econometrics*, 8, 15-30 (1993).

[28] Muller, N. & Yohai, V. Robust Estimates for ARCH Models Time Series Analysis, 23, 341-375 (2002).

[29] Eraker, B., Johannes, M. & Polson, N. The Impact of Jumps in Volatility and Returns. *Journal of Finance*. 58, 1269 & 1300 (2003).

[30] Mutunga, T. N., Islam, A. S. & Orawo, L. A. Estimation of Asymmetric GARCH Models: The Estimating Functions Approach. International Journal of Applied Science and Technology 4(5), 198 - 209 (2014).

[31] Godambe, V.P. & Thompson, M.E. An extension of the Quasi-likelihood Estimation, Journal of Statistical Planning and Inference, 22, 137-172 (1989).

[32] United Nations. Global Economic Outlook 2006. Global Economic and Monitoring Unit, Development Policy and Analysis Division, Department of Economic and Social Affairs, United Nations (2006).

[33] Pitt, M. K. & Shephard, N. Filtering via Simulation: Auxiliary Particle Filters. *Journal of the American Statistical Association*, 94, 590-599 (1999).

[34] Storvik, G. Particle Filters for State-Space Models with the Presence of Unknown Static *Parameters.* Institute of Electrical and Electronics Engineers (IEEE) Transactions on Signal Processing, *50, 281-289 (2002).*

[35] Johannes, M., Polson, N. & Stroud, J. Sequential Parameter Estimation in Stochastic Volatility Models with Jumps. Working paper, Graduate School of Business, University of Chicago
(2006).

[36] Raggi, D. & Bordignon, S. Sequential Monte Carlo Methods for Stochastic Volatility Methods with Jumps. Lecture Notes, University of Bologna/Padova, Italy, October 13, 2006, (2006).

[37] Zivot. E.V. Practical Issues in the Analysis of Univariate GARCH Models, Handbook of Financial Time Series, Springer, New York (2008).

[38] *Mandelbrot, B. The Variation of Certain Speculative Prices. Journal of Business, 36, 394-419 (1963).*