AMSE JOURNALS-2016-Series: Advances D; Vol. 21; N°1; pp 19-37
Submitted March 2016; Revised June 2, 2016, Accepted July 15, 2016

A Generalized t-Distribution based Filter for Stochastic Volatility Models

Onyeka-Ubaka, J. N.*, Abass, O.**, Okafor, R. O.*

*Department of Mathematics, University of Lagos, Lagos, Nigeria.
**Department of Computer Studies, Bells University, Ota, Nigeria.
(Email: jonyeka-Ubaka@unilag.edu.ng; okaforray@yahoo.com; olabass@unilag.edu.ng)

Abstract

The paper introduces a modified BL-GARCH (1, 1) model and the generalized student-t distribution with one skewness parameter and two tail parameters which offers 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. The paper evaluates the parameters of BL-GARCH (1, 1) and BL-GARCH (1, 1)-Volume model from Gaussian and non-Gaussian frameworks. Our empirical observation of asset returns shows that squared returns are positively autocorrelated; the reversion of volatility to the mean; they exhibit excess kurtosis (the fourth moment of returns), or fatter tails, relative to a normal distribution.

Key words: Squared returns, Gaussian, Non-Gaussian, Autocorrelation, Generalized t-distribution

1 Introduction

Volatility modeling plays a critical role in mathematical finance and statistical applications. The ability to estimate and forecast volatilities for different assets and groups of assets leads to a better understanding of current and future financial risk. The GARCH family models attempt to capture the autocorrelation of squared returns, the reversion of volatility to the mean, as well as the excess kurtosis. The first and simplest model is an ARCH model Engle [6], which stands for Autoregressive Conditional Heteroskedasticity. The AR comes from the fact that these models are autoregressive models in squared returns, while the conditional comes from the fact that in these models, next period's volatility is conditioned on information of the current period. Heteroskedasticity means non-constant volatility. The stock returns have 'heavy tails' or 'outliers prone' probability distributions. One reason for outliers may be that the conditional variance is not constant, and the outliers occur when the variance is large. The return on an asset is

$$r_t = \mu + z_t \sigma_t$$

where $z_t \overset{i.i.d.}{\sim} \mathrm{N}(0,1)$ the residual return at time t, $r_t - \mu$ as

$$\varepsilon_t(\theta) = z_t \sigma_t(\theta) \tag{1.1}$$

In an ARCH(1) model, next period's variance only depends on last period's squared residual so a crisis that caused a large residual would not have the sort of persistence that we observe after actual crises. This has led to an extension of the ARCH model to a Generalized ARCH (GARCH) model Bollerslev [2], which is similar in spirit to an ARMA model. In a GARCH (1, 1) model,

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Just like the ARCH model, the GARCH model can be extended to the GARCH (p, q) model as follows:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \qquad (1.2)$$

where $\alpha_0 > 0,\ \alpha_i \geq 0,\ i = 1,\ 2,\ \cdots,\ q \quad \beta_j \geq 0,\ j = 1,_2 2,\ \cdots,\ p.$ These conditions on parameters ensure strong positivity of the conditional variance (1.2). The model is covariance stationary if $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$. Its long-run average variance (unconditional variance) is equal to

$$\sigma^2 \equiv \varepsilon^2 \equiv \alpha_0 / (1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j) \quad . \qquad (1.3)$$

The exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model Nelson [19] has a somewhat different intellectual heritage but imply particular forms of conditional heteroskedasticity. This model attracts a significant attention because it can capture not only the asymmetric variance effect but also relax the non-negativity constraints imposed on the coefficients by GARCH model. The asymmetric models provide an explanation for the so called leverage effect, that is, an unexpected price drop increases volatility more than an analogous unexpected price increase. The EGARCH (p, q) model put forward by Nelson [20] provides a first explanation for the σ_t^2, that depends on both size and the sign of lagged residuals. He proposed the following form for the evolution of the conditional variance:

$$\ln(\sigma_t^2) = \omega + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2) + \sum_{i=1}^{q} \alpha_i [\phi\ Z_{t-i} + \gamma(|Z_{t-i}| - \mathrm{E}|Z_{t-i}|)] \qquad (1.4)$$

where $\varepsilon_t^2 = z_t^2 \sigma_t^2, \quad z_t \sim \mathrm{NI}D(0,\ 1)$, the parameters $\omega,\ \beta_j,\ \alpha_i$ are not restricted to be non-negative. The term $|z_{t-i}| - \mathrm{E}|z_{t-i}|$ is positive if the error term is larger than its expected value and negative otherwise. The parameter ϕ is the persistence of the volatility process that allows also for the volatility clustering feature. In empirical applications, this parameter is close to 1 even though it is assumed that $|\phi| < 1$. This condition implies that the conditional variance process is stationary, condition that is inherited by the returns (Su et al [28], Yu [30], Tsay and Ruey [29], Zhu and Wang [31] and Zhu [31]). This parameter is in general negative because high volatility induces expectations of lower future returns. However, the recent uncovering of the relation between volatility and trading volumes has led others such as Schwert [25], Gallant et al [8] as well as Jones et al [12] to propose an information-based variance model to explain the observed phenomenon. Najand and Yung [18], Locke and Sayers [13], Sharma et al [24], and Miyakoshi [15] showed that the current trading volume explain the persistence of variance represented by the ARCH-type effect.

Following earlier work of Storti and Vitale [26] and adapting Mohler [16] nonlinear representation of bilinear model, the state space representation of a bilinear model (of order m) in the control theory literatures is of the general form

$$\underline{x}(t) = \underline{A}\underline{x}(t-1) + \underline{B}\underline{\varepsilon}_t + \underline{C}\underline{x}(t-1)\underline{\varepsilon}_{t-1} \qquad (1.5)$$
$$X_t = \underline{H}\underline{x}(t)$$

where the system matrix $\underline{A}$ and the input matrix $\underline{B}$ are square matrices of order $(m \times m)$; the state vector $\underline{x}$ and the control vector $\underline{\varepsilon}$ are column vectors of order $(m \times 1)$. The input $\underline{\varepsilon}$ is a usually unobservable random process and the systems coefficient matrices are to be estimated.

If the paper nests the GARCH model and (1.5), the BL-GARCH model is given as

$$y_t = \mu + e_t \qquad (1.6)$$
$$\varepsilon_t^2 = z_t^2 \sigma_t^2$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{k=1}^{r} c_k \sigma_{t-k} \varepsilon_{t-k} \qquad (1.7)$$

Introducing the daily trading volume variable as additional information into the BL-GARCH model, we obtain a modified bilinear generalized autoregressive conditional heteroskedasticity-volume (BL-GARCH (p, q)-Volume) model given as

$$\varepsilon_t \big| (v_t,\ \varepsilon_{t-1},\ \varepsilon_{t-2},\ \cdots) \sim \mathrm{N}(0,\ \sigma_t^2)$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \gamma_i v_{t-1} + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{k=1}^{r} c_k \sigma_{t-k} \varepsilon_{t-k} \qquad (1.8)$$

where

$$\alpha_0 > 0,\ \ \alpha_i > 0\ \ i = 1,\ \cdots,\ q;\ \ \gamma_i > 0\ \ i = 1,\ \cdots,\ q;\ \ \beta_j > 0\ \ j = 1,\ \cdots,\ p\ \ and\ c_k^2 < 4\alpha_i\beta_j \qquad (1.9)$$

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's 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 and v_t is the daily trading volume, which is used as a proxy variable for the current information flow to the market.

If $\underline{c}$ = 0, the model (1.7) 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.

If $\gamma_i = 0$, the model (1.8) reduces to the state representation of the BL-GARCH (p, q) model. A positive γ_i suggests that the larger the information variable v_t, the larger the conditional variance. If the information variable (the trading volume variable) has explanatory power, γ_i is expected to be positive and at the same time, either α_i or β_j (or both) should become smaller and statistically insignificant.

2 Non-Gaussian Distributions with tail parameters

Let the innovations $(\varepsilon_t)_{t\in Z}$ have a conditional non-Gaussian:

(a) The Student-t distribution

$$f\{z_t(\theta);\eta\} = (2\pi)^{-1/2}(s^2)^{-1/2}\left(\frac{v}{2}\right)^{-1/2}\Gamma\left(\frac{v+1}{2}\right)\Gamma^{-1}\left(\frac{v}{2}\right)\left\{1+\frac{(x-\mu)^2}{vs^2}\right\}^{-(v+1)/2} \tag{2.1}$$

(b) The Generalized Student-t distribution

$$f\{z_t(\theta);\eta,q,\lambda_1,\lambda_2\} = \frac{\Gamma\left(\frac{v+1+iq}{2}\right)\left(\frac{v+1-iq}{2}\right)}{(\lambda_1+\lambda_2)2^{(1-v)}\pi(s)^{\frac{1}{2}}\Gamma(v)}\left(1+\frac{t^2}{v}\right)^{-(v+1)/2} \times \left(\lambda_1 \exp(q \ \ \arctan\frac{t}{v^{\frac{1}{2}}})\right) + \left(\lambda_2 \exp(-q \ \ \arctan\frac{t}{v^{\frac{1}{2}}})\right) \tag{2.2}$$

where

$-\infty < t < \infty$, q is a complex number, λ_1, $\lambda_2 \geq 0$. λ_1 and λ_2 are the left and right tail parameter respectively, v is the degrees of freedom. The standardized t deviate $t = (x-\mu)/s$ has distribution $t(0, 1, v)$, where x is the observations, μ is the mean and s is the standard deviation of the observations.

Note: $\lambda_1, \lambda_2 \geq 0$ is a necessary condition because the probability density function must always be positive. Also the normalization constant

$$\frac{\Gamma\left(\frac{v+1+iq}{2}\right)\left(\frac{v+1-iq}{2}\right)}{(\lambda_1+\lambda_2)2^{(1-v)}\pi(s)^{-\frac{1}{2}}\Gamma(v)}$$

of (2.2) is real, because the integrand is a real function on $(-\infty, \ \infty)$. It is clear that if $q = 0$ in (2.2), the usual Student-t distribution is derived. Moreover, for $q = 0$, the normalization constant of distribution (2.2) is equal to the normalization constant of Student-t distribution. The kurtosis of the Student-t distribution is

$$\mathrm{E}\left[\varepsilon_t^4\right] = \frac{3(v-2)}{v-4}$$

which is greater than three if $v < \infty$.

The MLE estimator $\hat{\theta}$ maximizes the log-likelihood function l_t given by

$$l_t = n\left[\log\Gamma\frac{(v+1)}{2} - \log\Gamma(\frac{v}{2}) - \frac{1}{2}\log\pi(v-2)\right] - \frac{1}{2}\sum_{t=1}^{n}\left\{\log(\sigma_t^2) + (v+1)\log[1+\frac{\varepsilon_t^2}{\sigma_t^2(v-2)}\right\} \qquad (2.3)$$

where $2 < v \le \infty$ and Γ is the Euler gamma function defined by $\Gamma(x) = \int_0^\infty t^{x-1}e^{-1}dt$. When $v \to \infty$, we have the normal distribution, so that the smaller the value of *v* the fatter the tails.

This means that for large *v*, the product

$$\left(\frac{v}{2}\right)^{-1/2}\Gamma\left(\frac{v+1}{2}\right)\Gamma^{-1}\left(\frac{v}{2}\right)$$

tends to unity, while the right-hand bracket in (2.1) tends to $e^{-(1/2\sigma^2)(x-\mu)^2}$.

The score function is given

$$\frac{\partial l_t}{\partial\theta} = \sum_{t=1}^{n}\left[\frac{v+1}{v-2}\frac{\varepsilon_t}{\sigma_t^2}\left(1+\frac{\varepsilon_t^2}{\sigma_t^2}\right)^{-1}\right]\frac{\partial\mu_t}{\partial\theta} + \frac{1}{2}\sum_{t=1}^{n}\left[\frac{v+1}{v-2}\frac{\varepsilon_t^2}{\sigma_t^2}\left(1+\frac{\varepsilon_t^2}{\sigma_t^2(v-2)}\right)^{-1} - 1\right]\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta} \qquad (2.4)$$

and the Hessian matrix is given by

$$\begin{aligned}\frac{\partial^2 l_t}{\partial\theta\partial\theta'} &= \frac{v+1}{v-2}\sum_{t=1}^{n}\left(1+\frac{\varepsilon_t^2}{(v-2)\sigma_t^2}\right)^{-1}\frac{\varepsilon_t}{\sigma_t^2}\left[\frac{2}{v-2}\left(1+\frac{\varepsilon_t^2}{(v-2)\sigma_t^2}\right)^{-1}\frac{\varepsilon_t}{\sigma_t^2} - 1\right]\frac{\partial\mu_t\partial\mu_t}{\partial\theta\partial\theta'}\\ &+ \frac{v+1}{v-2}\sum_{t=1}^{n}\left(1+\frac{\varepsilon_t^2}{(v-2)\sigma_t^2}\right)^{-1}\frac{\varepsilon_t}{\sigma_t^4}\left[\frac{1}{v-2}\left(1+\frac{\varepsilon_t^2}{(v-2)\sigma_t^2}\right)^{-1}\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right]\frac{\partial\mu_t}{\partial\theta}\frac{\partial\sigma_t^2}{\partial\theta'}\\ &+ \frac{v+1}{v-2}\sum_{t=1}^{n}\left(1+\frac{\varepsilon_t^2}{(v-2)\sigma_t^2}\right)^{-1}\frac{\varepsilon_t}{\sigma_t^4}\left[\frac{1}{v-2}\left(1+\frac{\varepsilon_t^2}{(v-2)\sigma_t^2}\right)^{-1}\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right]\frac{\partial\sigma_t^2}{\partial\theta}\frac{\partial\mu_t}{\partial\theta'}\\ &+ \frac{1}{2}\sum_{i=1}^{n}\frac{1}{\sigma_t^4}\left[1+\frac{v+1}{v-2}\frac{\varepsilon_t^2}{\sigma_t^2}\left(1+\frac{\varepsilon_t^2}{(v-2)\sigma_t^2}\right)^{-1}\right]\left[\frac{\varepsilon_t^2}{(v-2)\sigma_t^2}\left(1-\frac{\varepsilon_t^2}{(v-2)\sigma_t^2}\right)^{-1} - 2\right]\frac{\partial\sigma_t^2}{\partial\theta}\frac{\partial\sigma_t^2}{\partial\theta'}\end{aligned}$$

The generalized student-t distribution with one skewness parameter and two tail 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. This is because downward movement of the markets is followed by higher volatilities than upward movement of the same magnitude, see Muller and Yohai [17], Eraker et al [7], Gourieroux [9], Diongue et al [5], Onyeka-Ubaka [21]. So it is important to use BL-GARCH (1, 1) and BL-GARCH (1, 1)-Volume models to capture asymmetric shocks to volatility. This distribution function will be acceptable if it converges to the probability density of the standard normal distribution.

Preposition 2.1

If $\mathrm{T}(t, v, q, \lambda_1, \lambda_2)$ converges to the probability density of the standard normal distribution N(0, 1) as $v \to \infty$, that is

$$\lim_{v\to\infty}(t, v, q, \lambda_1, \lambda_2) = \mathrm{N}(t,\ 0,\ 1) \qquad (2.5)$$

then $\mathrm{T}(t, v, q, \lambda_1, \lambda_2)$ is valid for the generalized student-t distribution.

Proof

Taking limit,

$$\lim_{v\to\infty}\left(1+\frac{t^2}{v}\right)^{\frac{v+1}{2}}\left(\lambda_1\exp(q\ \arctan\frac{t}{v^{\frac{1}{2}}})\right)+\left(\lambda_2\exp(-q\ \arctan\frac{t}{v^{\frac{1}{2}}})\right)=\lambda_1+\lambda_2\exp\left(\frac{-t^2}{2}\right) \tag{2.6}$$

Using dominant convergence theorem (DCT) and noting that the DCT states that if for a continuous and integrable function g(x), we have

$$|f_v(x)\le g(x)|,$$

then

$$\lim_{v\to\infty}\int_a^b f_v(x)dx=\int_a^b \lim_{v\to\infty} f_v(x)dx \tag{2.7}$$

Considering the limit relation (2.6), we obtain

$$\lim_{v\to\infty}(t,v,q,\lambda_1,\lambda_2)=\frac{\lim_{v\to\infty}\left(1+\frac{t^2}{v}\right)^{-\frac{v+1}{2}}\left(\lambda_1\exp(q\ \arctan\frac{t}{v^{\frac{1}{2}}})+\lambda_2\exp(-q\ \arctan\frac{t}{v^{\frac{1}{2}}})\right)}{\int_{-\infty}^{\infty}\lim_{v\to\infty}\left(1+\frac{t^2}{2}\right)^{-\frac{v+1}{2}}\left(\lambda_1\exp(q\ \arctan\frac{t}{v^{\frac{1}{2}}})+\lambda_2\exp(-q\ \arctan\frac{t}{v^{\frac{1}{2}}})\right)dt}$$

$$=\frac{\lambda_1+\lambda_2\exp\left(-\frac{t^2}{2}\right)}{\int_{-\infty}^{\infty}\lambda_1+\lambda_2\exp\left(-\frac{t^2}{2}\right)dt} \tag{2.8}$$

$$=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}t^2\right)\ =\ \mathrm{N}(t,\ 0,\ 1) \tag{2.9}$$

This shows that the generalized student-t distribution converges to the normal distribution as the number of samples tends to infinity.
This completes the proof.

Preposition 2.2

Let the process $\{\varepsilon_t\}$ be defined by (1.7) and let BL-GARCH (1, 1) be given as

$$\sigma_t^2=\alpha_0+\alpha_1\varepsilon_{t-1}^2+\gamma_1 v_{t-1}+\beta_1\sigma_{t-1}^2+c_1\varepsilon_{t-1}\sigma_{t-1} \tag{2.10}$$

then $\{\varepsilon_t\}$ follows a distribution function.

Proof

From (1.7), it is obvious that the BL-GARCH (1, 1) processes are stationary if the process $\{\sigma_t^2\}$ is stationary. The weak stationarity condition of the BL-GARCH (1, 1) model is constructed using a multiplicative specification of the BL-GARCH (1, 1) model, which is also convenient for the construction of the quasi-strict stationarity conditions:

From (2.10), we have

$$\sigma_0\prod_{i=1}^{t}(\beta+\alpha z_{t-i}^2)+\alpha_0\left(1+\sum_{h=1}^{t-1}\prod_{i=1}^{h}(\beta+\alpha z_{t-i}^2)\right) \tag{2.11}$$

$$= \sigma_0 G_t + \alpha_0 \sum_{h=0}^{t-1} G_h \qquad (2.12)$$

where $z_t = \varepsilon_t / \sigma_t$, $G_k = G_{k-1}(\beta + \alpha z_{t-k}^2)$, G_0 = 1. The mean of the conditional variances in a BL-GARCH (1, 1) model can directly be calculated from equation (2.12). Assuming normality, one has

$$\mathrm{E}(G_h) = \mathrm{E}(G_{h-1}(\beta + \alpha z_{t-h}^2)) = \mathrm{E}(G_{h-1})\mathrm{E}(\beta + \alpha z_{t-h}^2) = (\alpha + \beta)^h \qquad (2.13)$$

So the unconditional variance becomes

$$\mathrm{E}(\sigma_t^2) = \alpha_0 \sum_{h=0}^{\infty} \mathrm{E}(G_h) = \alpha_0 \sum_{h=0}^{\infty} (\alpha + \beta)^h + \alpha z_{t-h}^2) = \frac{\alpha_0}{1 - \alpha - \beta} \qquad \alpha + \beta < 1$$

$$= \infty \qquad \alpha + \beta \geq 1 \qquad (2.14)$$

(This condition is similar to the one given by Bollerslev [2] for stationarity solution of the GARCH (1, 1) process. Note that the stationarity condition is independent of c_1 and γ_1.

To calculate the variance of the conditional variances of the BL-GARCH (1, 1) model we need to know the third- and fourth-order moments of the distributions. Once these moments are known, it is again straightforward to calculate these variances and covariances. Assuming normality, one can derive that

$$\mathrm{E}(G_h^2) = \mathrm{E}(G_{h-1}^2)\mathrm{E}(\beta + \alpha z_{t-1}^2)^2 = (\beta + 2\alpha\beta + 3\alpha^2)^h \qquad (2.15)$$

$$\mathrm{E}(G_h G_k) = \mathrm{E}(G_k^2 \prod_{i=1}^{h-k} (\beta + \alpha z_{t+i}^2)) = (\beta + \alpha)^{h-k} (\beta^2 + 2\alpha\beta + 3\alpha^2)^k \qquad (2.16)$$

$$\mathrm{E}(\sigma_t^2) = \alpha_0^2 \mathrm{E}\left(\sum_{h=0}^{\infty} G_h \right)^2 = \alpha_0^2 \left[\sum_{h=0}^{\infty} \mathrm{E}(G_h^2) + 2 \sum_{h=0}^{\infty} \sum_{k=0}^{h-1} \mathrm{E}(G_h G_k) \right]$$

$$= \frac{\alpha_0^2 (1 + \alpha + \beta)}{\left((1 - (\beta^2 + 2\alpha\beta + 3\alpha^2))(1 - (\alpha + \beta))\right)} \qquad (\beta^2 + 2\alpha\beta + 3\alpha^2) < 1 \qquad (2.17)$$

$$= \infty \qquad (\beta^2 + 2\alpha\beta + 3\alpha^2) \geq 1$$

Therefore, the variance of the conditional variances is given by

$$Var(\sigma_t^2) = \mathrm{E}(\sigma_t^2)^2 - (\mathrm{E}(\sigma_t^2))^2$$

$$= \frac{2\alpha_0^2 \alpha^2}{\left((1 - (\beta^2 + 2\alpha\beta + 3\alpha^2))(1 - (\alpha + \beta))^2\right)} \qquad (\beta^2 + 2\alpha\beta + 3\alpha^2) < 1$$

$$= \infty \qquad (\beta^2 + 2\alpha\beta + 3\alpha^2) \geq 1 \qquad (2.18)$$

In the case of more general error distributions such as student-*t* and generalized student-t, the mean and variance of σ_t^2 can be calculated by using the moments of a standardized t-distributed random variable, *z*:

$$\mathrm{E}(\sigma_t^2) = \alpha_0 \Big/ (1 - \alpha \mathrm{E}(z^2) - \beta) \qquad \alpha \mathrm{E}(z^2 + \beta) < 1 \qquad (2.19)$$

$$Var(\sigma_t^2) = \frac{\alpha_t^2[\mathrm{E}(z^4) - \mathrm{E}(z^2)]\alpha^2}{[(1-(\beta^2 + 2\mathrm{E}(z^2)\alpha\beta + \mathrm{E}(z^4)\alpha^2))(1-(\mathrm{E}(z^2)\alpha + \beta))^2]} \qquad \beta^2 + 2\mathrm{E}(z^2)\alpha\beta + \mathrm{E}(z^4)\alpha^2 < 1 \qquad (2.20\ )$$

The covariance of σ_t^2 and σ_h^2 can be calculated using equations (2.12) and (2.16):

$$= (\alpha\mathrm{E}(z^2) + \beta)^{t-h} Var(\sigma_t^2) \qquad t > h \qquad (2.21)$$

The fourth order moment of the process $\{\varepsilon_t\}$ exists if and only if

$$\beta_1^2 + c_1^2 + S\alpha_1^2 + 2\alpha_1\beta_1 < 1$$

where

$$S = \mathrm{E}(\varepsilon_t^4)$$

and it is equal to

$$\mathrm{E}\left(\varepsilon_t^4\right) = \frac{S\alpha_0^2(1+\alpha_1+\beta_1)}{(1-\alpha_1-\beta_1)(1-\beta_1^2 - c_1^2 - S\alpha_1^2 - 2\alpha_1\beta_1)} \qquad (2.22)$$

Let k be any finite positive integer. Then, when the distribution of $\{\varepsilon_t\}$ is generalized error, the second moment of $\{\varepsilon_t^{2k}\}$ is given by

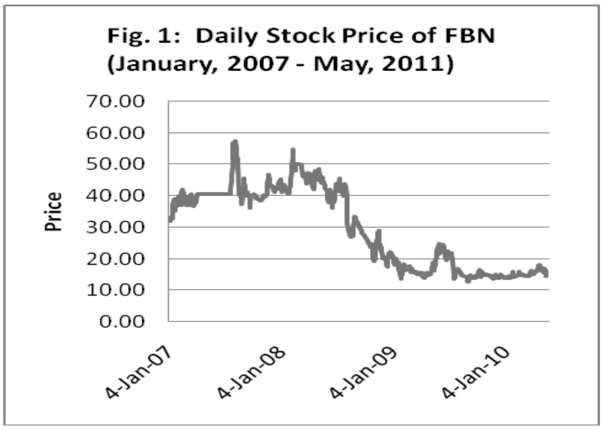
$$\mathrm{E}\left(\varepsilon_t^{4k}\right) = \left(\frac{2k\left[\omega - \gamma\Gamma\left(\frac{v+1+iq}{2}\right)\left(\frac{v+1-iq}{2}\right)\right]\sum_{i=1}^{q}\alpha_i}{(\lambda_1+\lambda_2)2^{(1-v)}\pi(s)^{-\frac{1}{2}}\Gamma(v)} \left(1+\frac{t^2}{v}\right)^{-(v+1)/2} \right) \mu_{4k}^{(g)} \mathrm{B}_{0,k}^{(g)} \times \prod_{j=1}^{p}\left(1-\beta_j\right)^{-1} \qquad (2.23)$$

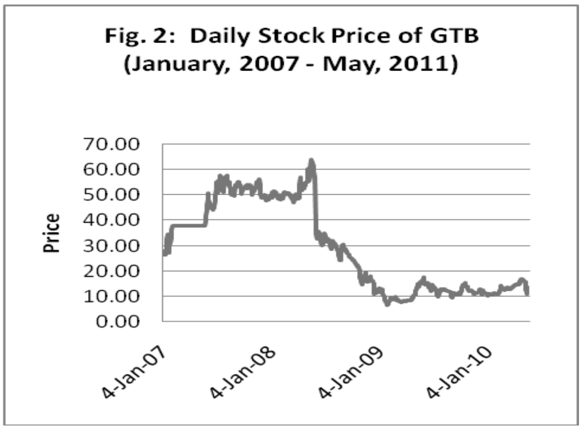
$$\times \left(\lambda_1 \exp(q \ \ \arctan\frac{t}{v^{\frac{1}{2}}}) \right) + \left(\lambda_2 \exp(-q \ \ \arctan\frac{t}{v^{\frac{1}{2}}} \right)$$

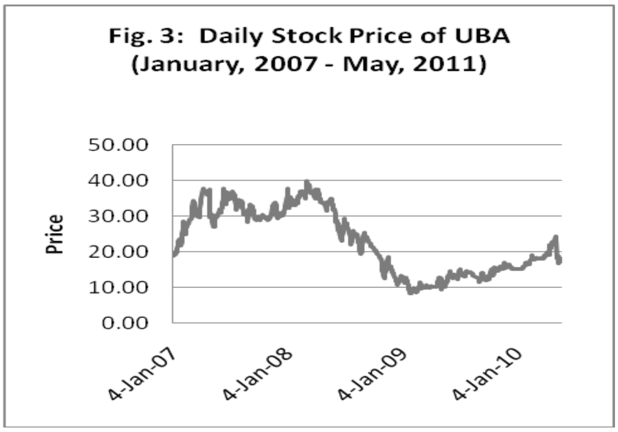
It follows that all the factors in (2.22) are positive so we conclude that the BL-GARCH (1, 1) process has the so-called leptokurtic distribution. If $(1-(\beta^2 + 2\mathrm{E}(z^2)\alpha\beta + \mathrm{E}(z^4)\alpha^2)) > 0$ and $v > 2$, the variances and covariances are finite and time independent. The BL-GARCH (1, 1) conditional variance process is covariance stationary. Hence, the proof.

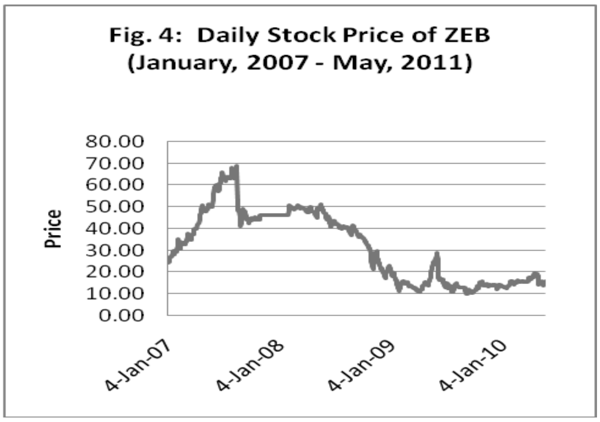
3 Empirical Study of Real Data

Figure 1, 2, 3 and 4 detail plots of the series (daily stock prices of selected banks on the floor of the Nigeria Stock Exchange (2007-2011). The data plotted display a non-stationary pattern with a decreasing trending behaviour with higher variability and lower level of the stock values toward the sample period for First Bank of Nigeria (FBN), Guaranty Trust Bank (GTB) and Zenith Bank (ZEB) while the plot for the United Bank for Africa (UBA) fluctuated, showed an upward trending behaviour toward the tail end of 2009 to 2010 and eventually fell in stock values at the end of the sample period.









The transformed data depict the continuously compounded daily returns of the selected banks. The plots show that the returns were more volatile over some time periods and became very volatile toward the end of the study period. This pattern of alternating quiet and volatile periods of substantial duration is referred to as volatility clustering. A visual inspection shows clearly, that the mean process for the different banks are not statistically significantly different from zero, but the variance changes over time, so the return series is not a sequence of independently and identically distributed (i.i.d.) random variables. A characteristic of asset returns, which is noticeable from the figures, is volatility clustering first noted by Mandelbrot [14]: “Large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes”.

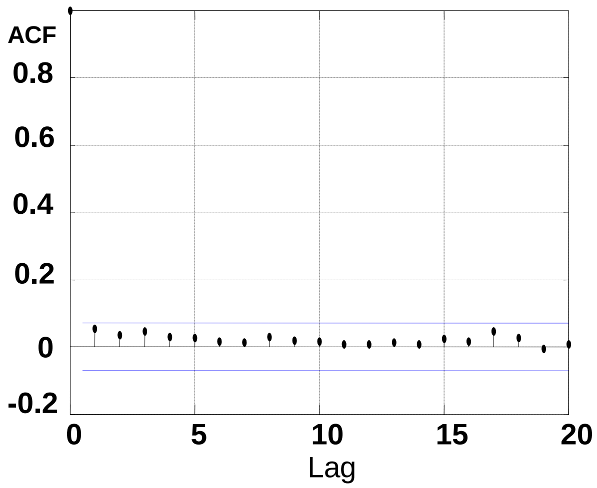


Fig. 5a: Sample ACF of FBN Squared Return Data

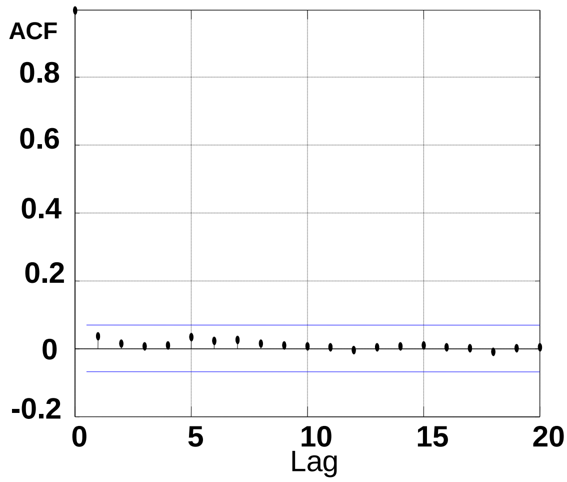


Fig. 5b: Sample ACF of GTB Squared Return Data

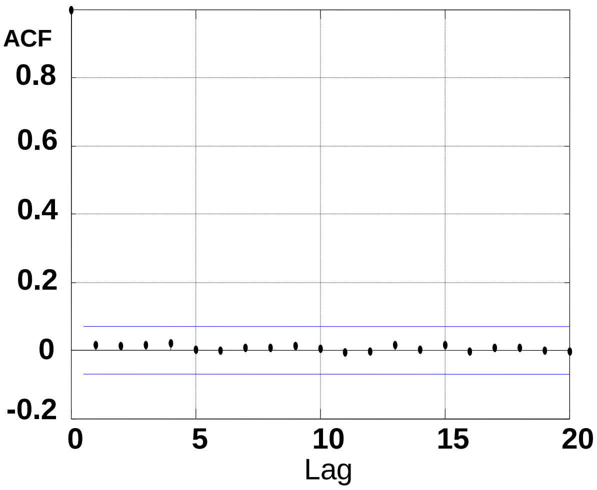


Fig. 5c: Sample ACF of UBA Squared Return Data

To select a suitable stochastic model, the study follows the three iterative steps of identification, estimation and diagnostic checks recommended by Box and Jenkins [3], Box et al [4]. For model identification, the study plotted the autocorrelation functions, the partial autocorrelation functions, the inverse autocorrelation functions and the inverse partial autocorrelation functions for the $y_t = \log p_t - \log p_{t-1}$. The volatility clustering observed in selected return data gives us a hint that they may not be independently and identically distributed, otherwise the variance would be constant. If the series values are truly independent, then nonlinear instantaneous transformations such as taking logarithms, absolute values, or squaring preserve independence Onyeka-Ubaka and Abass [22]. However, the same is not true of correlation, as correlation is only a measure of linear dependence. Higher-order serial dependence structure in data can be explored by studying the autocorrelation structure of the squared returns (of greater sampling variability but with more manageability in terms of statistical theory).

Although the ACF of the observed returns exhibits little autocorrelation, the ACF of the squared returns may still indicate significant correlation and persistence in the second-order moments. This is evident in the plots of the ACF of the squared returns in Figures 5a, b, c and d below.

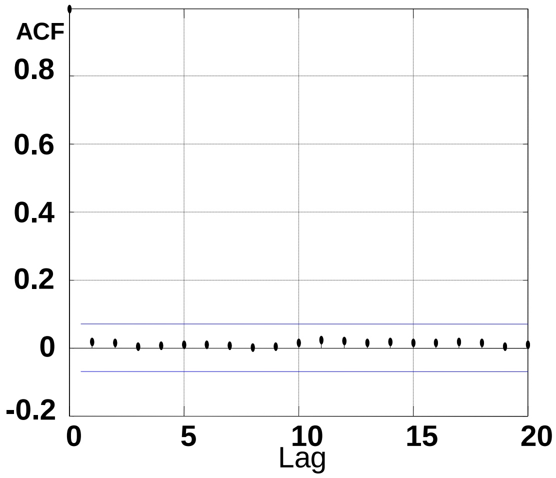


Fig. 5d: Sample ACF of ZEB Squared Return Data

The returns in Figure 5 show substantial evidence of GARCH effects (heteroskedasticity) as judged by the autocorrelations of the squared residuals. The first order autocorrelations of FBN and GTB are 0.093 and 0.085, and they gradually decline to 0.023 and 0.011 respectively after 10 lags confirming that the fall in the prices of stock causes some shareholders to sell their shares before maturity while those of UBA fluctuates around zero and the ZEB gradually increase after 10 lags. The increase after lag 10 evinced that price volatility persists for a long period drawing the bank stocks to have leverages. However, these autocorrelations are not large, but they are positive and very significant. This is an implication of the BL-GARCH (1, 1) model. Quantifying the preceding qualitative checks for correlation using formal hypothesis tests, such as the Ljung-Box-Pierce Q-test and Engle's ARCH test, both functions return identical outputs. The test p-values shown in the third column are all zero, resounding rejecting the "no ARCH" hypothesis. The first output, H, is a Boolean decision flag. H = 0 implies that no significant correlation exists (that is, do not reject the null hypothesis). H = 1 means that significant correlation exists (that is, reject the null hypothesis). The remaining outputs are the p-value (pValue), the test statistic (Stat) and the critical value of the Chi-square distribution (CriticalValue). The results show that there is significant serial correlation in the squared returns when the researcher tests them with the same inputs. The results of the two tests and diagnostics estimates are summarized in Tables 1, 2 and 3 below:

Table 1: Ljung-Box-Pierce Q-test

Lag	H	PValue	Stat	CriticalValue
10	0.9991	0	1.4309	18.3070
15	1.0000	0	2.1450	24.9958
20	1.0000	0	3.6852	31.4104

Table 2: Engle's ARCH test

Lag	H	PValue	Stat	CriticalValue
10	0.9991	0	1.4389	18.3070
15	0.9999	0	2.2313	24.9958
20	1.0000	0	3.6407	31.4104

Table 3: Diagnostics Estimates of the Selected Banks BL-GARCH (1, 1) Model

Bank	Jarque-Bera	P-Value (Chi^2)	Skewness	Kurtosis	Log-lik	AIC	BIC
FBN Normal	2876.1695	0.0000	-1.1861	11.8227	1835.17	2531.90	2542.81
Student-t	2957.4381	0.0000	-2.1485	11.3241	1839.40	2070.86	2037.11
Gen. Student-t	2976.1109	0.0000	-2.1306	10.7983	1839.28	2147.39	2137.04
GTB Normal	6806.3915	0.0000	-1.8134	16.5783	1831.09	2651.24	2673.41
Student-t	6813.8694	0.0000	-1.6518	21.6132	1841.76	2348.59	2318.67
Gen. Student-t	6813.8209	0.0000	-1.7502	21.8794	1841.25	2316.46	2358.27
UBA Normal	2857.1063	0.0000	-2.9160	31.1992	1829.23	3542.70	3560.13
Student-t	2897.4356	0.0000	-3.7821	33.2867	1832.69	3576.87	3592.10
Gen. Student-t	2897.3722	0.0000	-3.6735	33.2794	1832.14	3572.09	3592.03
ZEB Normal	4034.3057	0.0000	-3.1875	36.6163	1815.87	4631.30	4640.47
Student-t	4159.5378	0.0000	-3.4917	35.7580	1822.94	4628.19	4634.89
Gen. Student-t	4159.5256	0.0000	-3.4907	35.0812	1822.01	4625.07	4634.66
BL-GARCH(1,1)-V Gen. Student-t	2918.3766	0.0000	-2.5051	20.1601	1822.72	8345.07	8729.46

These tests show significant evidence in support of GARCH effects. Each of these extracts the sample mean from the actual returns. This is consistent with the definition of the conditional mean equation in which the innovations process is $\varepsilon_t = y_t - \mu_t$, and μ_t is the mean of y_t. Evidence of time dependence is found using Ljung-Box statistics, which is robust to heteroskedasticity and reported for autocorrelations up to 20 lags. The statistics show strong serial correlations in both levels of the return series. This is consistent with the results of Storvik [27], Johannes et al [11], Raggi and Bordignon [23], who found that serial correlations in DJIA returns are significant but unstable and depend on the sample period.

A non- constant variance of asset returns should lead to a non-normal distribution. Figure 6 represents the histogram and the student-t distribution of the stock market prices of the selected banks. Non-stationarity in the conditional variances is not the only possible source of non-stationarity for stock return series. The level of stock prices and trading volume may also be non-stationary. The volatile behaviour of the disturbances contradicts the assumption of normally distributed disturbances. The disturbances are therefore modelled with the BL-GARCH (1, 1) using generalized student-t distributed random variables with unknown common degrees of freedom. The increase in the variance does not occur when the disturbances are generalized student-t distributed because with a lower value of the parameter λ one can also explain occurrence of several rather large values of the disturbances, that is, heteroskedasticity.

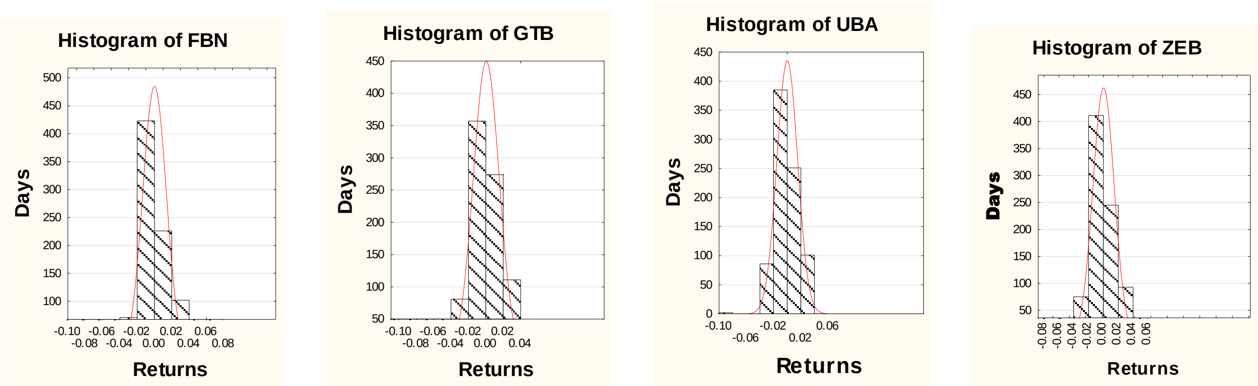


Fig. 6: Histogram and Student-t Distribution of the Selected Banks

Figure 6 shows evidence of fat tails since the kurtosis is positive and evidence of skewness, which means that the tails are either heavier or lighter than the usual student-t distribution. Looking at the plot of GTB stock prices, the student-t distribution tends to infinity. This prompts the study to use the generalized student-t distribution to capture the extreme values. Kurtosis and the narrow bands in plot are hints of conditional heteroskedasticity. These models, although able to capture the leptokurtosis, could not account for the existence of non-linear temporal dependence as the volatility clustering observed from the data. The Jarque-Bera (JB) test decisively rejects the normal distribution (see Table 3). The JB test is a test statistic for testing whether the series is normally distributed. The test statistic measures the difference of the skewness and kurtosis of the series with those from the normal distribution.

$$\text{Jarque-Bera} = \frac{N-k}{6}\left(S^2 + \frac{(K-3)^2}{4}\right)$$

where S is the skewness, K is the kurtosis, and *k* represents the number of estimated coefficients used to create the series. Under the null hypothesis of a normal distribution, the JB statistic is distributed as χ^2 with 2 degrees of freedom.

The data was estimated by methods of the Maximum Likelihood Estimator (MLE) using MATLAB (R2008b) soft ware. The parameter estimates are presented in Tables 4, 5 and 6 below:

Table 4: Conditional Variance GARCH (1, 1) Model Parameter Estimation Results

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{v}$
Gaussian	**-0.2189 0.03281* (0.01973, -11.0948)**	**0.1732 0.01055* (0.00924, 18.7446)**	**0.9215 0.00683* (0.01387 66.4384)**	**-**
Student-t	**-0.2803 0.03218* (0.03041, -9.2174)**	**0.3805 0.04163* (0.0479, 7.9436)**	**0.9391 0.05061* (0.05713, 16.4379)**	**3.2496 0.00934* (0.1106 29.3816)**
Gaussian	**$1.8145e^{-005}$ 0.02756***	**0.1750 0.00385***	**0.8243 0.00194***	**-**

FBN	(2.4118e^{06}, 7.5234)	(0.01721, 10.1685)	(0.01371, 60.1240)	
Student-t FBN	2e^{-007} 0.03471* (5.2309e^{-008}, 0.8234)	0.3634 0.00317* (0.04007, 9.0691)	0.6365 0.00513* (0.01941, 32.7924)	2.3682 0.02716* (0.10355, 22.8701)
Gaussian GTB	0.0006 0.03538* (6.7353e^{-005}, 8.5820)	0.0642 0.01730* (0.07362, 0.8720)	0.5510 0.06719* (0.04693, 11.7409)	-
Student-t GTB	0.0001 0.05983* (3.027e^{-005}, -3.3940)	0.3621 0.04752* (0.07485, 4.8377)	0.5951 0.07683* (0.06528, 9.1161)	6.2527 0.05423* (0.69364, 9.0143)
Gaussian UBA	0.0002 0.03795* (1.950e^{-005}, -7.9737)	0.2924 0.00975* (0.04097, 7.1369)	0.7065 0.01709* (0.03558, 19.8567)	-
Student-t UBA	2e^{-007} 0.04189* (5.2013e^{-008}, 3.8452)	0.3799 0.01493* (0.04163, 9.1256)	0.6201 0.00873* (0.02267, 27.3533)	4.7919 0.03891* (0.28051, 17.0821)
Gaussian ZEB	2.194e^{-005} 0.05147* (2.1776^{e-006}, 10.0754)	0.2285 0.01756* (0.01069, 5.4888)	0.7714 0.01642* (0.00890, 86.6742)	-
Student-t ZEB	2e-007 0.04297* (5.9104^{e-008}, 3.3838)	0.3081 0.00798* (0.03615, 8.5228)	0.6918 0.01587* (0.02081, 33.2436)	4.6334 0.03476* (0.31416, 14.7485)

Table 5: Conditional Variance BL-GARCH (1, 1) Model Parameter Estimation Results

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{c}_1$	$\hat{v}$	$\hat{\lambda}_1$	$\hat{\lambda}_1$
Gaussian	**-0.2950 0.01842*** **(0.01292, 22.4458)**	**0.2832 0.00352*** **(0.00823, 34.4107)**	**0.9511 0.01065*** **(0.01937 49.1017)**	**-0.0118 0.00537*** **(0.00568, -2.0775)**	**-**		
Student-t	**-0.3002 0.02178*** **(0.05049, -5.9457)**	**0.4051 0.00413*** **(0.02279, 17.7753)**	**0.9238 0.00683*** **(0.07014, 13.1708)**	**-0.1036 0.01075*** **(0.01773, -5.8432)**	**3.3429 0.03937*** **(0.15028 22.2443)**		
G. Std-t	**-0.3106 0.00572*** **(0.04512, -6.8839)**	**0.3916 0.00294*** **(0.02187, 17.9058)**	**0.9162 0.00158*** **(0.05218 17.5585)**	**-0.5314 0.02416*** **(0.08023, -6.6235)**	**3.5841 0.01927*** **(0.17209, 20.8269)**	**20.3**	**17.1**
Gaussian FBN	**1.8145e^{-005} 0.02481*** **(2.4118e^{06}, 7.5234)**	**0.17507 0.00322*** **(0.017201, 10.1778)**	**0.82493 0.00723*** **(0.013713, 60.1560)**	**0.00591 0.00537*** **(0.05316, 0.1112)**	**-**		
Student-t FBN	**2e^{-007} 0.02178*** **(5.2309e^{-008}, 0.8234)**	**0.3634 0.00318*** **(0.04007, 9.0691)**	**0.6765 0.00659*** **(0.01941, 34.8531)**	**-0.0856 0.01075*** **(0.01575, -5.4349)**	**2.0682 0.05137*** **(0.1035, 19.9826)**		
G. Std-t FBN	**2e-005 0.03217*** **(5.1308e^{-008}, 0.8234)**	**0.3831 0.00594*** **(0.04107, 9.3279)**	**0.6483 0.06253*** **(0.02394, 27.0802)**	**-0.0880 0.01075*** **(0.01455, -6.0481)**	**2.5618 0.05137*** **(0.1065, 24.05446)**	**24.4**	**10.8**
Gaussian GTB	**-3.7376 0.02338*** **(0.69453, -5.3815)**	**0.7452 0.00813*** **(0.04746, 15.0695)**	**0.4548 0.05247*** **(0.09867, 4.6103)**	**-0.0760 0.05327*** **(0.03834, -1.9823)**	**-**		
Student-t GTB	**-1.1341 0.02057*** **(0.29667, -3.8226)**	**0.4904 0.05164*** **(0.07519, 6.5221)**	**0.8391 0.06945*** **(0.04146, 2.0239)**	**-0.0283 0.06197*** **(0.05117, -5.5306)**	**2.2948 0.36092*** **(0.74276, 3.0896)**		
G. Std-t GTB	**-3.1893 0.02178*** **(0.36262, -8.7952)**	**0.5190 0.01006*** **(0.05761, 9.0089)**	**0.4398 0.05352*** **(0.01567, 28.0664)**	**-0.0980 0.05169*** **(0.05212, -1.8803)**	**7.8974 0.06673*** **(0.84215, 9.3778)**	**18.4**	**7.8**
Gaussian UBA	**-1.8027 0.04358*** **(0.13484, -13.3690)**	**1.0215 0.01062*** **(0.05546, 18.4191)**	**0.7333 0.01965*** **(0.01753, 41.8311)**	**-0.1560 0.00951*** **(0.03737, -4.1745)**	**-**		
Student-t UBA	**-0.4415 0.03761*** **(0.09842, -4.4859)**	**0.4915 0.00923*** **(0.04615, 10.6501)**	**0.94166 0.00714*** **(0.01326, 71.0078)**	**-0.1234 0.00714*** **(0.03633, -3.3966)**	**2.1625 0.03637*** **(0.31338, 6.9006)**		
G. Std-t UBA	**-0.6134 0.02516*** **(0.18793, -3.2640)**	**0.4257 0.00623*** **(0.06183, 6.8850)**	**0.8419 0.00412*** **(0.03128, 26.9150)**	**-0.1760 0.01398*** **(0.02904, -6.0606)**	**5.9763 0.03637*** **(0.52160, 11.4576)**	**20.9**	**14.0**
Gaussian ZEB	**2.194e^{-005} 0.05329*** **(2.1776^{e-006}, 10.0754)**	**0.2285 0.01067*** **(0.01069, 21.3751)**	**0.7717 0.01594*** **(0.00894, 86.3199)**	**-0.0294 0.01553*** **(0.03008, -0.9774)**	**-**		
Student-t ZEB	**2e^{-007} 0.04087*** **(5.9104^{e-008}, 3.3838)**	**0.3081 0.00771*** **(0.03615, 8.5228)**	**0.6918 0.01439*** **(0.02081, 33.2436)**	**-0.0189 0.03369*** **(0.02644, -0.7148)**	**3.6334 0.05213*** **(0.31416, 11.5654)**		
G. Std-t ZEB	**2e^{-005} 0.02106*** **(5.8104^{e-008}, 4.3927)**	**0.3176 0.01194*** **(0.025809, 12.3058)**	**0.6543 0.00339*** **(0.02167, 30.1938)**	**-0.0297 0.03369*** **(0.01962, -1.5138)**	**5.6735 0.00168*** **(0.40267, 14.0897)**	**24.1**	**13.5**

Table 6: Conditional Variance BL-GARCH (1, 1)-Volume Model Parameter Estimation Results

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{c}_1$	$\hat{v}$	$\hat{\gamma}$

Gaussian	**-0.2858 0.01482*** **(0.01290, 22.1550)**	**0.2532 0.00235*** **(0.00852, 29.7183)**	**0.7312 0.01060*** **(0.01937 37.7491)**	**-0.0187 0.00157*** **(0.00569, -3.3393)**	**-**	
Student-t	**-0.2902 0.02117*** **(0.05091, -5.7003)**	**0.3815 0.00514*** **(0.02973, 12.8322)**	**0.6526 0.00782*** **(0.00719, 90.7650)**	**-0.1033 0.01170*** **(0.01746, -5.9164)**	**4.3429 0.03639*** **(0.56028 8.6373)**	**0.2853**
Gaussian FBN	**1.9145e-^{005} 0.02081*** **(2.4018e^{06}, 7.5234)**	**0.1751 0.00332*** **(0.01901, 9.2109)**	**0.8543 0.00423*** **(0.01371, 62.3122)**	**-0.0951 0.00532*** **(0.06345, -1.4988)**	**-**	
Student-t FBN	**2e^{-007} 0.02079*** **(5.2309e^{-008}, 0.8234)**	**0.3019 0.00372*** **(0.0410, 7.3634)**	**0.6167 0.00635*** **(0.01915, 32.2037)**	**-0.0872 0.01093*** **(0.01575, -5.5365)**	**2.3872 0.04187*** **(0.10355, 23.0536)**	**-0.1697**
Gaussian GTB	**-3.7376 0.02208*** **(0.60534, -6.1579)**	**0.6723 0.06101*** **(0.04746, 14.1656)**	**0.3148 0.05042*** **(0.08653, 3.6380)**	**-0.0763 0.05207*** **(0.03845, -1.9844)**	**-**	
Student-t GTB	**-1.3415 0.02152*** **(0.26796, -5.0063)**	**0.6398 0.06715*** **(0.07198, 8.8886)**	**0.3809 0.06105*** **(0.04156, 9.1651)**	**-0.0820 0.06170*** **(0.05172, -1.5855)**	**6.2948 0.30012*** **(0.74276, 8.4749)**	**0.5242**
Gaussian UBA	**-1.8027 0.03548*** **(0.15328, -11.7608)**	**0.2215 0.01072*** **(0.09147, 2.4216)**	**0.8203 0.01826*** **(0.07952, 10.3156)**	**-0.1605 0.00981*** **(0.06713, -2.3909)**	**-**	
Student-t UBA	**-0.4417 0.02571*** **(0.09842, -4.4879)**	**0.4125 0.05823*** **(0.06154, 6.7030)**	**0.8260 0.05018*** **(0.07613, 10.8499)**	**-0.1432 0.04063*** **(0.07633, -1.8761)**	**3.6125 0.03117*** **(0.3133, 11.5304)**	**0.2389**
Gaussian ZEB	**2.204e^{-005} 0.05409*** **(2.1806^{e-006}, 10.0754)**	**0.2285 0.01167*** **(0.03198, 7.1451)**	**0.7714 0.01395*** **(0.02899, 26.6092)**	**-0.02319 0.01652*** **(0.05138 -0.4513)**	**-**	
Student-t ZEB	**2e^{-006} 0.04182*** **(5.8903^{e-008}, 2.8398)**	**0.3082 0.00817*** **(0.06158, 5.0049)**	**0.6738 0.01039*** **(0.05218, 12.3130)**	**-0.0189 0.02136*** **(0.02644 -0.7148)**	**4.8231 0.01273*** **(0.36432, 13.2386)**	**0.0956**

where

The asterisks (*) are the P-values of the estimated parameters.

The values in parenthesis, say (a, b), are the standard errors and t-statistics respectively.

Tables 4, 5 and 6 represent conditional variance GARCH (1, 1)**,** BL-GARCH (1, 1) and BL-GARCH (1, 1)-Volume model parameter estimation results respectively. Results reveal that parameter estimates are satisfactory in that the standard errors are small and the t-statistic for GARCH parameters is high. It is clear from the analysis that estimate $\hat{\alpha}_1$ and $\hat{\beta}_1$ in the BL-GARCH (1, 1) and BL-GARCH (1, 1)-Volume model are significant at the 5% level with the volatility coefficient greater in magnitude. Hence, the hypothesis of constant variance is rejected, at least within sample. Furthermore, the stationarity condition is satisfied for the three distributions, as $\hat{\alpha}_1 + \hat{\beta}_1 < 1$ at the maximum of the respective log-likelihood functions. The estimated asymmetric volatility response $\hat{c}_1$ is negative and significant for all models except for GTB confirming the usual expectation in stock markets where downward movements (falling returns) are followed by higher volatility than upward movements (increasing returns). The results also follow the empirical findings of Storti and Vitale [26], in that the kurtosis strongly depends on the leverage-effect response parameter. The results for statistics indicate that the BL-GARCH (1, 1) as well as the GARCH (1, 1) processes is appropriate for modelling the conditional variance of the selected banks return. However, the goodness-of-fit statistics as well as the residuals diagnostics indicate that the BL-GARCH (1, 1) performs better in describing the conditional variance of the selected banks return. Moreover, the possible usefulness of using fat-tailed innovations for the GARCH (1, 1) and BL-GARCH (1, 1) models seem to be confirmed by the log-likelihood values and the AIC in Table 3. Using Akaike [1], the BL-GARCH (1, 1) model with minimum AIC was selected as the best.

The BL-GARCH (1, 1) conditional variance model that best fits the observed data is

$$\sigma_t^2 = -0.3106 + 0.3916\varepsilon_{t-1}^2 + 0.9162\sigma_{t-1}^2 - 0.5314\varepsilon_{t-1}\sigma_{t-1}$$

where

$$\hat{\alpha}_0 = -0.3106, \quad \hat{\alpha}_1 = ARCH(1) = 0.3916, \quad \hat{\beta}_1 = GARCH(1) = 0.9162$$

$$and \quad c_1 = leverage\,effect = -0.5314$$

The model for individual bank estimates are given as

$$\sigma_t^2 = 2e^{-007} + 0.3634\varepsilon_{t-1}^2 + 0.6765\sigma_{t-1}^2 - 0.0856\varepsilon_{t-1}\sigma_{t-1} \quad \text{......... FBN}$$

$$\sigma_t^2 = -3.1893 + 0.5190\varepsilon_{t-1}^2 + 0.4398\sigma_{t-1}^2 + 0.0980\varepsilon_{t-1}\sigma_{t-1} \quad \text{......... GTB}$$

$$\sigma_t^2 = -0.6134 + 0.4257\varepsilon_{t-1}^2 + 0.8419\sigma_{t-1}^2 - 0.1760\varepsilon_{t-1}\sigma_{t-1} \quad \text{......... UBA}$$

$$\sigma_t^2 = 2e^{-005} + 0.3176\varepsilon_{t-1}^2 + 0.6543\sigma_{t-1}^2 - 0.0297\varepsilon_{t-1}\sigma_{t-1} \quad \text{.......... ZEB}$$

From the results obtained, the BL-GARCH (1, 1) model with generalized student-t distribution fits GTB, UBA and ZEB data better while the First Bank of Nigeria data follows the student-t BL-GARCH (1, 1) model. This is because adding more parameters in modelling the FBN data does not improve the parameter estimates of the FBN. The parameter λ is therefore a good approximation of the degree up to which one is able to explain the variance/kurtosis of the disturbances. The generalized student-t distribution of λ for BL-GARCH conditional variances lies almost completely above 2.0 such that the conditional variances of the disturbances are finite. In Table 6, the coefficients for trading volume variable, γ is positive, greater than 0 and significant at 5% level for GTB, UBA and ZEB. This analysis is suggesting that trading volume increases due to good news in the market. Coefficients γ shows that a positive impact of volume on stock returns also generate less impact on volatility of the market. Thus, the inclusion of a trading volume variable in the variance process accounts for some of the observed GARCH persistence and asymmetric effect embedded in the volatility of returns for the sampled period. This analysis also shows that the three models: GARCH (1, 1), BL-GARCH (1, 1) and most importantly the BL-GARCH (1, 1)-Volume model explicitly established that the recent news of trading volume can be used to improve the prediction of stock price volatility in Nigeria banking sector and by extension any banks in the world.

Conclusion

This paper establishes: the bilinear generalized autoregressive conditional heteroskedasticity (BL-GARCH) model capture empirical characteristics present in high frequency financial time series data; the evaluation of parameters of BL-GARCH (1, 1) and BL-GARCH (1, 1)-Volume model from

Gaussian and non-Gaussian frameworks; leptokurtosis in banks' stock returns. That is, the kurtosis of the selected banks' stock returns exceeds the kurtosis of a standard Gaussian distribution, showing marginal distributions with heavier tails and thin centres; and strong volatility, persistence and asymmetry of the selected banks, so the inclusion of contemporaneous trading volume in the BL-GARCH model results in a positive relationship between trading volume and volatility. The trading volume affects the flow of information, supporting the validity of the mixture of distributions hypothesis (MDH). The MDH provides an explanation for volatility and volume by linking changes in price, volume and the rate of information flow. The asymmetric effect of bad news on volatility is higher when contemporaneous trading volume is included, although market shocks, whether positive or negative, have similar effects on conditional volatility. Thus, we conclude that trading volume is a useful tool for predicting the volatility dynamics of the selected banks, and by extension other mega banks, in Nigeria.

References

[1] Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 1974; AC-19: 716-723.

[2] Bollerslev T. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 1986; 31: 307-327.

[3] Box G.E.P., Jenkins G. W. Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, 1976.

[4] Box G. E. P., Jenkins G. M., Reinsel G. C. *Time Series Analysis, Forecasting and Control,* Prentice-Hall, Englewood Cliffs, 1994.

[5] 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*, 2010; 7: 4-10.

[6] Engle R. F. Autoregressive Conditional Heteroskedasticity with Estimates of United Kingdom Inflation. *Econometrica*, 1982; 50: 987-1007.

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

[8] Gallant R., Rossi P. E., Tauchen G. Stock Prices and Volume. *Review of Financial Studies*, 1992; 5: 199-242.

[9] Gourieroux C. ARCH Models and Financial Applications, Springer Series in Statistics, New York, 2007; 67-72.

[10] Hall P. and Yao Q. Inference in ARCH and GARCH Models with Heavy-Tailed Errors. *Econometrica*, 2003; 71: 285-317.

[11] 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.

[12] Jones C. G., Kaul G. K., Lipson, M. L. Transactions, Volumes and Volatility. *Review of Financial Studies* 1994; 7: 631-651.

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

[14] Mandelbrot B. The Variation of Certain Speculative Prices. *Journal of Business*, 1963; 36: 394-419.

[15] Miyakoshi T. ARCH versus Information-Based Variances: Evidence from the Tokyo Stock Market. *Japan and the World Economy*, 2002; 14: 215-231.

[16] Mohler R. R. Bilinear control processes – with applications to engineering, ecology and medicine. Academic Press, New York, 1973.

[17] Muller N., Yohai V. Robust Estimates for ARCH Models Time Series Analysis, 2002; 23: 341-375.

[18] Najand M., Yung K. A GARCH Examination of Relationship between Volume and Price Variability in Futures Markets. *Journal of Futures Markets*, 1991; 11: 613-621.

[19] Nelson D. B. Stationarity and Persistence in the GARCH (1, 1) Model, *Econometric Theory*, 1990; 6: 318-334.

[20] Nelson D.B. Conditional Heteroskedasticity in assets returns: A new approach, the modelling of financial time series, *Econometrica*, 1991; 59: 347-370.

[21] Onyeka-Ubaka J. N. A Modifiled BL-GARCH Models for Distributions with Heavy Tails. A Ph.D Thesis, University of Lagos, Nigeria, 2013.

[22] Onyeka-Ubaka J. N., Abass O. Central Bank of Nigeria (CBN) and the Future of Stocks in the Banking Sector. *American Journal of Mathematics and Statistics*,2013; 3(6): 407-416.

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

[24] Sharma J.L., Mougoue M., Kamath R. Heteroskedasticity in Stock Market Indicator Return Data: Volume versus GARCH Effects. *Applied Financial Economics*, 1996; 6: 337-342.

[25] Schwarz G. Estimating the dimension of a model. *Annals of Statistics*. 1978; 6: 461-464.

[26] Storti G., Vitale C. BL-GARCH models and asymmetries in volatility. *Statistical Methods and Applications.* 2003; 12: 19–40.

[27] Storvik G. (2002). Particle Filters for State-Space Models with the Presence of Unknown Static Parameters. *IEEE Transactions on Signal Processing*, 2002; 50: 281-289.

[28] Su Y., Yip Y., Wong, R. The Impact of Government Intervention on Stock Returns Evidence from Hong Kong. *International Review of Economics and Finance*, 2002; 11: 277-297.

[29] Tsay, Ruey Analysis of Financial Time Series, 2nd Edition, John Wiley & Sons Inc. New York, 2005; 97-153.

[30] Yu J. On Leverage in a Stochastic Volatility Model.*Journal of Econometrics*, 2005; 127: 165-178.

[31] Zhu F., Wang, D. A mixture integer-valued ARCH model. *Journal of Statistical Planning and inference*, 2010; 140: 2025-36.

[32] Zhu F. A negative binomial integer-valued GARCH model. *Journal of Time Series Analysis*, 2011; 32: 54-67.