

Heavy Tails Estimation in Nonlinear Models

Onyeka-Ubaka J. N.^{1,*}, Abass O.², Okafor R. O.¹

¹Department of Mathematics, University of Lagos, Akoka, Lagos, +234, Nigeria

²Department of Computer Science, University of Lagos, Akoka, Lagos, +234, Nigeria

Abstract A generalized student t distribution technique based on estimation of bilinear generalized autoregressive conditional heteroskedasticity (BL-GARCH) model is described. The paper investigates from empirical perspective, among other things, aspects related to the economic and financial risk management and to its impact on volatility forecasting. The purposive sampling technique was applied to select four banks (First Bank of Nigeria (FBN), Guaranty Trust Bank (GTB), United Bank for Africa (UBA) and Zenith Bank (ZEB)) daily stock prices, considered to be more susceptible to volatility than other banks within the sampled period (January, 2007- May, 2011). The data collected were analyzed using MATLAB R2008b Software. The results show that the newly introduced generalized student-t distribution is the most general of all the useful distributions applied in the BL-GARCH model parameter estimation. They serve as general distributions for obtaining empirical characteristics such as volatility clustering, leptokurtosis and leverage effect between returns and conditional variances as well as capturing heavier and lighter tails in high frequency financial time series data.

Keywords BL-GARCH, Leverage effect, Leptokurtosis, Heteroskedasticity, Nonlinear

1. Introduction

Time varying parameter models have a long history in statistics. Modelling unequal variances in nonlinear time series is a challenging task. The ideas and techniques of modelling time series in these diverse areas of science and other related disciplines are proved to be useful and innovative by Box and Jenkins[4], Engle[9], Gilks and Berzuini[16] and[30]. Bilinear Generalized Autoregressive Conditional Heteroskedasticity (BL-GARCH) model is one of the major tools that Economists use to model financial markets' behaviour in the presence of political disorders, economic crises, wars or natural disasters. In such stress periods, prices of financial assets tend to fluctuate very profusely[30].

Statistically speaking, if the conditional variance of Y_t given Y_{t-1}, Y_{t-2}, \dots that is, $Var(Y_t | Y_{t-1}, Y_{t-2}, \dots)$ of a time series is not constant over time then the process Y_t is conditionally heteroskedastic. Heteroskedasticity refers to the random errors having unequal variances. In particular, a heteroskedastic model has $Cov(e) = Diag(\sigma_t^2)$, Christensen[7]. Let $Y = (Y_1, Y_2, \dots, Y_n)$ be a

Gaussian vector with mean vector μ and variance matrix Σ .

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}$$

If the expected value of all error terms when squared is the same at any given point, then the vector is homogeneous (homoskedastic) that is, $\sigma_1 = \sigma_2 = \dots = \sigma_n$. When this assumption does not hold, the vector is heteroskedastic, see Box, Jenkins and Reinsel[5], Bollerslev, Engle and Nelson[3]. There are several approaches to dealing with heteroskedasticity. If the error variance at different times is known, weighted regression is a good method. If as is the case with financial time series, the error variance is unknown, and must be estimated from the data, we can model the changing error variance with the bilinear generalized autoregressive conditional heteroskedasticity model.

2. Literature Review

In the past decades, many researchers have introduced informal and ad-hoc procedures to take account of the changes in the variance. One of the first authors to address variance changing over time was Mandelbrot (1963), who used recursive estimates of the variance for modelling

* Corresponding author:
 jonyeka-Ubaka@unilag.edu.ng (Onyeka-Ubaka J. N.)
 Published online at <http://journal.sapub.org/statistics>
 Copyright © 2014 Scientific & Academic Publishing. All Rights Reserved

volatility. Klein (1977) used rolling estimates of quadratic residuals.

Engle (1982) proposes Autoregressive Conditional Heteroskedasticity (ARCH (q)) models that seem to capture the empirical characteristics in the financial time series. The models have non-constant variances conditioned on the past, which is a linear aggregate of recent past disturbances. This means that the more recent news will be the fundamental information that is relevant for modelling the present volatility. Some leading scholars who studied ARCH models include: Geweke[14, 15], Diebold and Nerlove (1989), Engle[10], Jones, Kaul and Lipson[19] and Gouriéroux[17]. Harvey, Ruiz and Sentana (1992) established the existence of a few common factors explaining exchange rate volatility movements. Engle, Ng and Rothschild[11] show that U. S. bond volatility changes are closely linked across maturities. This commonality of volatility changes holds not only across assets within a market, but also across different markets. For example, Schwert[31] established that U. S. stock and bond volatilities move together, while Engle and Susmel[12], Hamao, Masulia and Ng[18] discovered close links between volatility changes across international stock markets. Lamoureux and Lastrapes[21] deduced that conditional heteroskedasticity may be caused by time dependence in the rate of information arrival to the market. They use the daily trading volume of stock markets as a proxy for such information arrival and confirm its significance. Mizrahi[25] associates ARCH models with the errors of the economic agent's learning processes. Cai[6] proposed the switching ARCH or SWARCH model in which there are several different ARCH models and that the economy switches from one to another following a Markov chain. In this model there can be extremely high volatility process which is responsible for events such as the stock market crash in 2009. That volatilities move together should be encouraging to model builders, since it indicates that a few common factors may explain much of the temporal variation in the conditional variances and covariances of asset returns.

Storti and Vitale[32] proposed BL-GARCH model in Gaussian framework. Diongue, Guegan and Wolff[9] extended their works using elliptical noise to capture the leverage effect or negative correlation between asset returns and volatility.

3. Methodology

The paper adapts recommended three iterative steps of Box-Jenkins to select a suitable stochastic model, Box and Jenkins (4), Box, Jenkins and Reinsel[5]. The steps are

- (i) Identification
- (ii) Estimation
- (iii) Diagnostic Checks

The aim of the identification stage is to determine the transformation required to produce stationarity and also the order of AR and MA operators for the ε_t series. In a typical

BL-GARCH modelling application, it is preferable that there are a minimum of about 80 data points in the ε_t series in order to get reasonable MLEs for the parameters. This identification starts with time series plot which may reveal one of the following characteristics:

- (i) trends either in the mean level or variance of the time series
- (ii) extreme values and outliers
- (iii) seasonality

At the estimation stage, estimates are usually calculated for the conditional mean, ARCH, GARCH and leverage effect parameters using Maximum Likelihood Estimates (MLE). In this paper, natural logarithms are applied to obtain the transformation required to produce stationarity. The normality assumption of the residuals is usually not critical for obtaining good parameter estimates. As long as the ε_t 's are independent and possess finite variance, reasonable estimates (Gaussian estimates) of the parameters can be obtained, Abass[1], Onyeka-Ubaka, Abass and Okafor[27]. Having observed that the conditional variance depends on the data, the paper adapted maximum likelihood method which is consistent and asymptotically normal. This is because financial time series data, for which BL-GARCH models are usually capable of capturing the characteristics, generate high frequency sampling of data.

The paper uses the Gaussian (Normal) and the non-Gaussian (Generalized Student t) distributions to allow the model fit both the tails and the central part of the conditional distribution present in high frequency financial time series data. The elliptical normalized distributions: the Normal and the Generalized Student t considered in this paper belong to exponential class. This is because their distributions can be expressed as

$$f(x; \theta) = \begin{cases} \exp[P(\theta)B(x) + C(x) + D(\theta)] \\ 0 \end{cases} \quad \text{elsewhere}$$

This implies that they have complete sufficient statistics for the estimators.

(a) The normal distribution is uniquely determined by its first two moments (Dallah, Okafor and Abass[8]). Hence, only the conditional mean and variance parameters enter the log-likelihood function

$$L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n \left\{ \log \left(\sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2} \right) \right\}, \quad (1)$$

where n is the sample size.

To obtain an analytical or numerical solution of the MLE, it is necessary to obtain the first derivative $\frac{\partial L(\theta)}{\partial \theta}$ and

solve $\frac{\partial L(\theta)}{\partial \theta} = 0$. Assuming $\psi = \theta$, we have the score functions as

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{t=1}^n \frac{\varepsilon_t}{\sigma_t^2} \frac{\partial \mu_t}{\partial \theta} + \frac{1}{2} \sum_{t=1}^n \frac{1}{\sigma_t^2} \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial \sigma_t^2}{\partial \theta} \quad (2)$$

Now, if the paper assumes that the innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ have a conditional non-Gaussian:

(b) The Student t distribution is given as

$$f[z_t(\theta); \eta] = (2\pi)^{-1/2} (s^2)^{-1/2} \left(\frac{v}{2} \right)^{-1/2} \Gamma\left(\frac{v+1}{2} \right) \Gamma^{-1}\left(\frac{v}{2} \right) \left\{ 1 + \frac{(x-\mu)^2}{vs^2} \right\}^{-(v+1/2)} \quad (3)$$

(c) The Generalized Student t distribution is given as

$$\begin{aligned} f\{z_t(\theta); \eta, q, \lambda_1, \lambda_2\} &= \frac{\Gamma\left(\frac{v+1+iq}{2} \right) \Gamma\left(\frac{v+1-iq}{2} \right)}{(\lambda_1 + \lambda_2) 2^{(1-v)} \pi(s)^{-1/2} \Gamma(v)} \left(1 + \frac{t^2}{v} \right)^{-(v+1/2)} \\ &\times \left[\lambda_1 \exp\left(q \arctan \frac{t}{v^{1/2}} \right) \right] + \left[\lambda_2 \exp\left(-q \arctan \frac{t}{v^{1/2}} \right) \right] \end{aligned} \quad (4)$$

where $-\infty < t < \infty$, q is a complex number. $\lambda_1, \lambda_2 \geq 0$. λ_1 and λ_2 are the left and right tail parameter respectively, v is the degrees of freedom. The standardized 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

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

of (4) is real, because the integrand is a real function on $(-\infty, \infty)$. It is clear that if $q = 0$ in (4), the usual Student t distribution is derived. Moreover, for $q = 0$, the normalization constant of distribution (4) is equal to the normalization constant of Student t distribution. The kurtosis of the Student t distribution is

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

which is greater than three if $v < 4$.

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

$$\begin{aligned} l_t &= n \left[\log \Gamma\left(\frac{v+1}{2} \right) - \log \Gamma\left(\frac{v}{2} \right) - \frac{1}{2} \log \pi(v-2) \right] \\ &\quad - \frac{1}{2} \sum_{t=1}^n \left\{ \log(\sigma_t^2) + (v+1) \log \left[1 + \frac{\varepsilon_t^2}{\sigma_t^2(v-2)} \right] \right\} \end{aligned} \quad (5)$$

where $2 < v \leq \infty$ and Γ is the Euler gamma function defined by $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. When $v \rightarrow \infty$, 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 (3) tends to $e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$.

The score function is given by

$$\frac{\partial l_t}{\partial \theta} = \sum_{t=1}^n \left[\left(\frac{v+1}{v-2} \right) \frac{\varepsilon_t}{\sigma_t^2} \left(1 + \frac{\varepsilon_t^2}{\sigma_t^2} \right)^{-1} \right] \frac{\partial \mu_t}{\partial \theta} + \frac{1}{2} \sum_{t=1}^n \left[\left(\frac{v+1}{v-2} \right) \frac{\varepsilon_t^2}{\sigma_t^2} \left(1 + \frac{\varepsilon_t^2}{\sigma_t^2(v-2)} \right)^{-1} - 1 \right] \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} \quad (6)$$

and the Hessian matrix is given by

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \theta \partial \theta'} &= \frac{v+1}{v-2} \sum_{t=1}^n \left(1 + \frac{\varepsilon_t^2}{(v-2)\sigma_t^2} \right)^{-1} \frac{\varepsilon_t}{\sigma_t^2} \left[\frac{2}{v-2} \left(1 + \frac{\varepsilon_t^2}{(v-2)\sigma_t^2} \right)^{-1} \frac{\varepsilon_t}{\sigma_t^2} - 1 \right] \frac{\partial \mu \partial \mu}{\partial \theta \partial \theta'} \\ &+ \frac{v+1}{v-2} \sum_{t=1}^n \left(1 + \frac{\varepsilon_t^2}{(v-2)\sigma_t^2} \right)^{-1} \frac{\varepsilon_t}{\sigma_t^2} \left[\frac{1}{v-2} \left(1 + \frac{\varepsilon_t^2}{(v-2)\sigma_t^2} \right)^{-1} \frac{\varepsilon_t}{\sigma_t^2} - 1 \right] \frac{\partial \mu}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'} \\ &+ \frac{v+1}{v-2} \sum_{t=1}^n \left(1 + \frac{\varepsilon_t^2}{(v-2)\sigma_t^2} \right)^{-1} \frac{\varepsilon_t}{\sigma_t^2} \left[\frac{1}{v-2} \left(1 + \frac{\varepsilon_t^2}{(v-2)\sigma_t^2} \right)^{-1} \frac{\varepsilon_t}{\sigma_t^2} - 1 \right] \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \mu_t}{\partial \theta'} \\ &+ \frac{1}{2} \sum_{t=1}^n \frac{1}{\sigma_t^4} \left[1 + \frac{v+1}{v-2} \frac{\varepsilon_t^2}{\sigma_t^2} \left(1 + \frac{\varepsilon_t^2}{(v-2)\sigma_t^2} \right)^{-1} \right] \left[\frac{\varepsilon_t^2}{(v-2)\sigma_t^2} \left(1 - \frac{\varepsilon_t^2}{(v-2)\sigma_t^2} \right)^{-1} - 2 \right] \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'} \end{aligned} \quad (7)$$

4. Results and Discussion

In this paper, a BL-GARCH model is introduced in which the innovation is composed of several sources of errors where each of the error sources has heteroskedastic specifications of the GARCH form. Since the error components cannot be separately observed given past observations, the independent variables in the variance equations are not measurable with respect to the available information set, which complicates procedures. Following earlier work of 2and Vitale[31] and adapting Mohler[26] 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

$$\underline{x}(t) = \underline{A} \underline{x}(t-1) + \underline{B} \underline{\varepsilon}_t + \underline{C} \underline{x}(t-1) \underline{\varepsilon}_{t-1} \quad (10)$$

$$\underline{X}_t = \underline{H} \underline{x}(t)$$

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 (10), the BL-GARCH model is given as

$$y_t = \mu + \varepsilon_t \quad (11)$$

$$\varepsilon_t^2 = \sigma_t^2 z_t^2 \quad (12)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{k=1}^r c_k \sigma_{t-k} \varepsilon_{t-k} \quad (13)$$

Given the parameter space, Ω , and the expected parameter vector, $\omega = (\alpha, \theta)$, the assumption is that the parameter $\omega_0 \in \Omega \subseteq R^{k+l+p+q+r+l}$ is in the interior of Ω , a compact parameter space.

Specifically, for any vector $\omega \in \Omega$, we assume that

(a) The AR and MA polynomials have no common roots and that all their roots lie outside the unit circle.

(b) $\alpha_0 > 0$, $\alpha_1, \dots, \alpha_q \geq 0$ and $\beta_1, \dots, \beta_p \geq 0$

(c) $c_i^2 < 4\alpha_i\beta_i$, for $i = 1 \dots, r$

(d) $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$

where the compact space is given as

$$\Theta = \left\{ \theta : \mu_1 \leq \mu \leq \mu_d, 0 < \omega_1 \leq \omega \leq \omega_d, 0 < \alpha_1 \leq \alpha \leq \alpha_d, 0 < \beta_1 \leq \beta \leq \beta_d \right\} \\ \subseteq \left\{ \theta : E[\ln(\beta + az^2)] < 0 \right\}$$

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 Pagan and Schwert[29], Locke and Sayers (1993), Linton[22], Muller and Yohai (2002), Eraker, Johannes and Polson[13]. So it is important to use BL-GARCH (1, 1) model to capture asymmetric shocks to volatility. This distribution function will be acceptable if it converges to the probability density of the standard normal distribution.

Proposition 4.1

If $f\{z_t(\theta), \eta, q, \lambda_1, \lambda_2\}$ as in (4) is distribution flexible, then it contains Pearson subordinate distributions.

Proof

The generalized student t distribution can be derived from a generalization of the Pearson differential equation as follows

$$\frac{df}{dz} = \frac{-g(z)f(z)}{h(z)} \quad (14)$$

where $g(z)$ and $h(z)$ are polynomials in the random variable z , and $f(z)$ is the density function of z . In the standard Pearson system, $g(z)$ is a polynomial in z of the degree less than or equal to one, whereas $h(z)$ is a polynomial in z of degree less than or equal to two. The generalized solution of (14) is

$$f(z) = \exp \left[- \int^z \left(\frac{g(s)}{h(s)} ds - \eta \right) \right], \quad z \in D \quad (15)$$

where the normalizing constant is given by

$$\eta = \ln \int \exp \left[- \int^z \left(\frac{g(s)}{h(s)} ds \right) \right] dz \quad (16)$$

The domain D of $f(z)$ in (16) is the open interval where $h(z)$ is positive.

The choice of $g(\cdot)$ and $h(\cdot)$ for the generalized student t distribution are

$$g(z) = \sum_{i=0}^{M-1} \alpha_i z^i \quad (17)$$

$$h(z) = \gamma^2 + z^2 \quad (18)$$

By substituting (17) and (18) in (14), the generalized student t distribution is given by

$$f(z) = \exp \left[\lambda_1 \tan^{-1} \left(\frac{z}{\gamma} \right) + \lambda_2 \ln(\gamma^2 + z^2) + \sum_{i=3}^M \lambda_i z^{i-2} - \eta \right] \quad -\infty < z < \infty \quad (19)$$

where

$$\eta = \log \int \exp \left[\lambda_1 \tan^{-1} \left(\frac{z}{\gamma} \right) + \lambda_2 \ln(\gamma^2 + z^2) + \sum_{i=3}^M \lambda_i z^{i-2} \right] dz \quad (20)$$

And the distribution parameters λ_i are functions of the parameters

$$\{\alpha_i, \quad i = 0, 1, \dots, M-1; \gamma\}$$

given in (17) and (18). Provided that in (19), $\lambda_M < 0$, all moments of the distribution exist. This distribution can exhibit a range of shapes including fat tails; sharp peaks and even multimodality (see Lye et al (1998)).

As the generalized student t distribution given by (20) is derived from an extension of the Pearson exponential family, it directly contains many of the Pearson subordinate distributions as special cases. In particular, from the point of view of the existing ARCH models, these special cases include the Normal and Student t distributions. The standard normal distribution occurs when $\lambda_3 = -0.5$ and all remaining parameters are zero. The Student t distribution occurs when $\lambda_2 = -0.5(1 + \gamma^2)$ and all remaining parameters are zero.

A special case which turns out to be important in the empirical application is given by (20) with $\lambda_i = 0$, $i > 2$ and $\lambda_2 = -0.5(1 + \gamma^2)$.

$$f(z) = \exp \left[\lambda_1 \tan^{-1} \left(\frac{z}{\gamma} \right) - 0.5(1 + \gamma^2) \ln(\gamma^2 + z^2) + \sum_{i=3}^M \lambda_i z^{i-2} - \tilde{\eta} \right] \quad -\infty < z < \infty \quad (21)$$

where $\tilde{\eta}$ is the normalized constant. This distribution is referred to as the skewed Student t distribution where skewness is controlled by the parameter, λ_1 . When $\lambda_1 = 0$, there is no skewness and the distribution becomes Student t distribution.

This completes the proof.

4.1. Empirical Study of Real Data

Four banks: First Bank of Nigeria (FBN), Guaranty Trust Bank (GTB), United Bank for Africa (UBA) and Zenith Bank (ZEB) were selected for the study by means of purposive sampling technique. 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 Central Bank of Nigeria (CBN) in August, 2009. Their volume of stocks traded on the floor of the Nigeria Stock Exchange (NSE) for the sample period were collected and analyzed using BL-GARCH model to observe the volatile nature of stocks within the sampled period. The series plots of the banks that exhibited heavier tails are shown below.

Figure 1 shows that the generalized student t distribution seems to be a more appropriate distribution for the selected banks data. The generalized student t distribution allows for situations where the tails are heavier or lighter than the usual student t distribution. The extreme values (left (λ_1) and right tails (λ_2)) for different banks are estimated automatically together with the plot of the P-P plot of the selected banks in Figure 2.

The probability-probability plots of the four banks show that they are primarily a few large outliers that cause the departures of the system from normality as obtainable from Figure 2. These departures are pointing out that there are other factors that interrupt the expected volatility of stock market prices of these banks on the floor of the Nigeria Stock Exchange. The factors may include among other things, giving loans to private and corporate firms to buy shares

without due process, lack of strategic management and regular supervision. If the residuals are normally distributed, the P-P plot should lie on a straight line.

The parameter results estimated by methods of the Maximum Likelihood Estimator (MLE) using MATLAB (R2008b) soft ware were given in Table 1.

Table 1 represents conditional variance BL-GARCH (1, 1) model parameter estimation results. Results reveal that parameter estimates are satisfactory (asymptotically unbiased, efficient and consistent) in that the standard errors are small and the t-statistic for GARCH parameters (β) is high. It is clear from the analysis that estimate $\hat{\alpha}_1$ and $\hat{\beta}_1$ in the BL-GARCH (1, 1) 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. Even when $\hat{\alpha}_1 + \hat{\beta}_1 > 1$, so long as $E[\ln(\alpha_1 z_t^2 + \beta_1)] < 0$, covariance stationarity is established. The estimated asymmetric volatility response \hat{c}_1 is negative and significant for all models 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[32], in that the kurtosis strongly depends on the leverage-effect response parameter. The results indicate that the BL-GARCH (1, 1) processes are appropriate for modelling the conditional variance of the selected banks return. Using Akaike (1974), 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

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

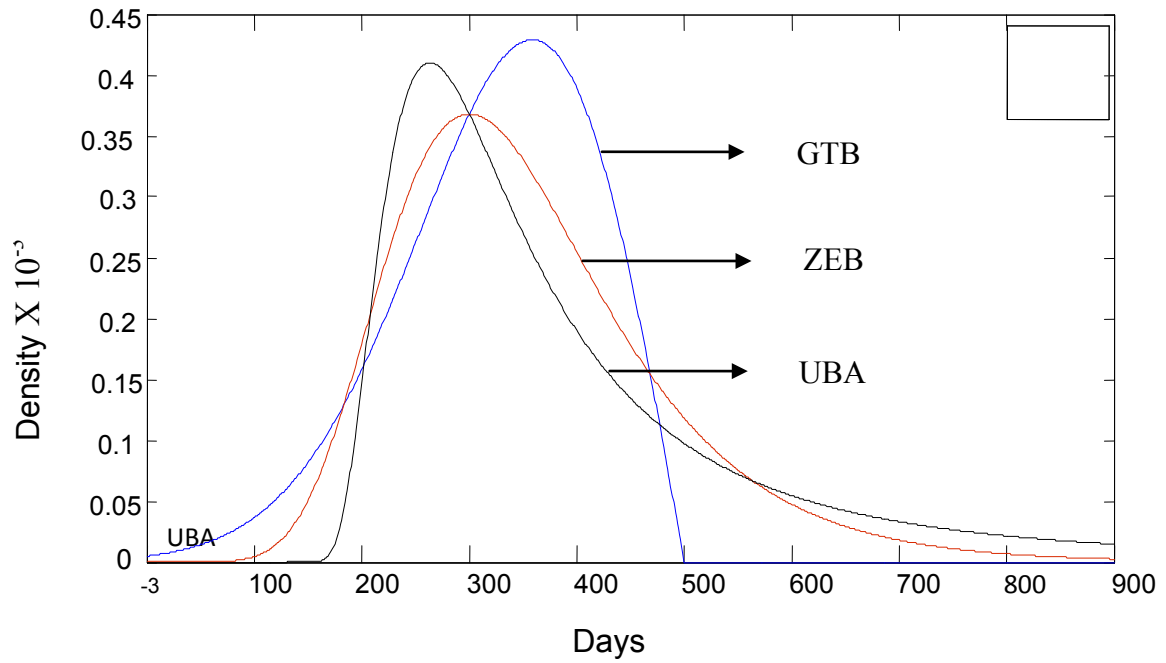


Figure 1. Plots of Generalized Student t Distribution for Banks that exhibit Heavier and Lighter Tails

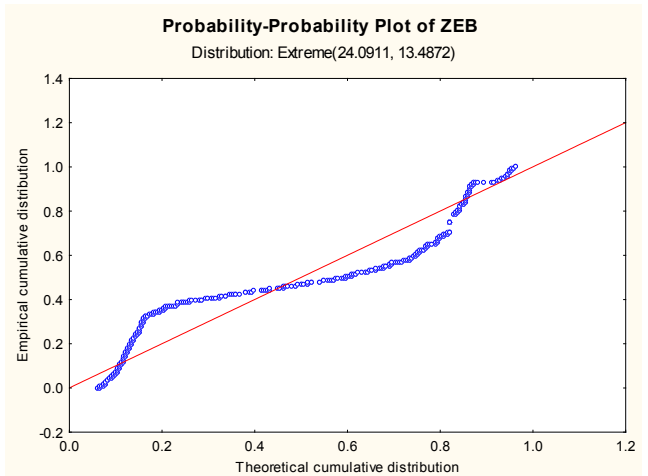
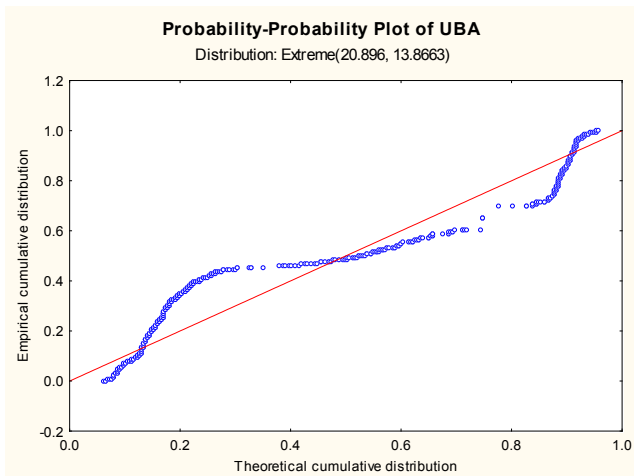
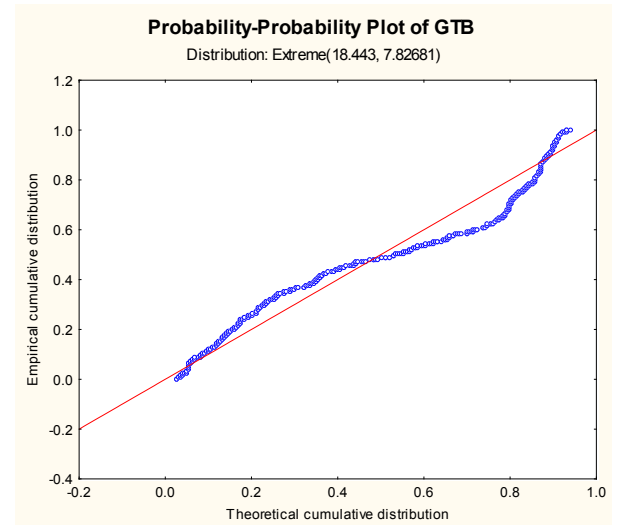
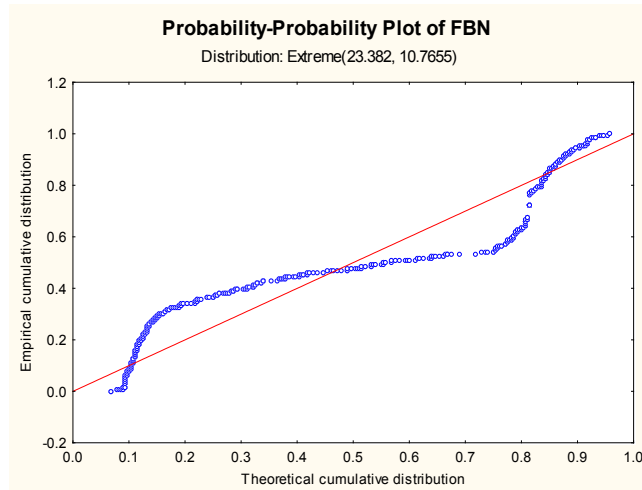


Figure 2. Probability-Probability (P-P) Plots with Extreme Values for the Four Banks

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

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	\hat{c}_1	\hat{v}	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Gaussian	-0.2950 (0.01292, 22.4458)	0.2832 (0.00823, 34.4107)	0.9511 (0.01937, 49.1017)	-0.0118 (0.00568, -2.0775)	-		
Student <i>t</i>	-0.3002 (0.05049, -5.9457)	0.4051 (0.02279, 17.7753)	0.9238 (0.07014, 13.1708)	-0.1036 (0.01773, -5.8432)	3.3429 (0.15028, 22.2443)		
G. Std <i>t</i>	-0.3106 (0.04512, -6.8839)	0.3916 (0.02187, 17.9058)	0.9162 (0.05218, 17.5585)	-0.5314 (0.08023, -6.6235)	3.5841 (0.17209, 20.8269)	20.3	17.1
Gaussian FBN	1.8145e ⁻⁰⁰⁵ (2.4118e ⁻⁰⁶ , 7.5234)	0.17507 (0.017201, 10.1778)	0.82493 (0.013713, 60.1560)	0.00591 (0.05316, 0.1112)	-		
Student <i>t</i> FBN	2e ⁻⁰⁰⁷ (5.2309e ⁻⁰⁰⁸ , 0.8234)	0.3634 (0.04007, 9.0691)	0.6765 (0.01941, 34.8531)	-0.0856 (0.01575, -5.4349)	2.0682 (0.1035, 19.9826)		
G. Std <i>t</i> FBN	2e ⁻⁰⁰⁵ (5.1308e ⁻⁰⁰⁸ , 0.8234)	0.3831 (0.04107, 9.3279)	0.6483 (0.02394, 27.0802)	-0.0880 (0.01455, -6.0481)	2.5618 (0.1065, 24.05446)	23.4	10.8
Gaussian GTB	-3.7376 (0.69453, -5.3815)	0.7452 (0.04746, 15.0695)	0.4548 (0.09867, 4.6103)	-0.0760 (0.03834, -1.9823)	-		
Student <i>t</i> GTB	-1.1341 (0.29667, -3.8226)	0.4904 (0.07519, 6.5221)	0.8391 (0.04146, 2.0239)	-0.0283 (0.05117, -5.5306)	2.2948 (0.74276, 3.0896)		
G. Std <i>t</i> GTB	-3.1893 (0.36262, -8.7952)	0.5190 (0.05761, 9.0089)	0.4398 (0.01567, 28.0664)	-0.0980 (0.05212, -1.8803)	7.8974 (0.84215, 9.3778)	18.4	7.8
Gaussian UBA	-1.8027 (0.13484, -13.3690)	1.0215 (0.05546, 18.4191)	0.7333 (0.01753, 41.8311)	-0.1560 (0.03737, -4.1745)	-		
Student <i>t</i> UBA	-0.4415 (0.09842, -4.4859)	0.4915 (0.04615, 10.6501)	0.94166 (0.01326, 71.0078)	-0.1234 (0.03633, -3.3966)	2.1625 (0.31338, 6.9006)		
G. Std <i>t</i> UBA	-0.6134 (0.18793, -3.2640)	0.4257 (0.06183, 6.8850)	0.8419 (0.03128, 26.9150)	-0.1760 (0.02904, -6.0606)	5.9763 (0.52160, 11.4576)	20.9	14.0
Gaussian ZEB	2.194e ⁻⁰⁰⁵ (2.1776e ⁻⁰⁰⁶ , 10.0754)	0.2285 (0.01069, 21.3751)	0.7717 (0.00894, 86.3199)	-0.0294 (0.03008, -0.9774)			
Student <i>t</i> ZEB	2e ⁻⁰⁰⁷ (5.9104e ⁻⁰⁰⁸ , 3.3838)	0.3081 (0.03615, 8.5228)	0.6918 (0.02081, 33.2436)	-0.0189 (0.02644, -0.7148)	3.6334 (0.31416, 11.5654)		
G. Std <i>t</i> ZEB	2e ⁻⁰⁰⁵ (5.8104e ⁻⁰⁰⁸ , 4.3927)	0.3176 (0.025809, 12.3058)	0.6543 (0.02167, 30.1938)	-0.0297 (0.01962, -1.5138)	5.6735 (0.40267, 14.0897)	24.1	13.5

where

The asterisks (*) are the P-values.

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

where

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

$$\hat{\beta}_1 = GARCH(1) = 0.9162$$

$$\text{and } \hat{c}_1 = \text{leverage effect} = -0.5314$$

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) models. 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 GTB, UBA and ZEB series confirm these statements as seen in Figure 1.

5. Conclusions

The newly introduced generalized student t distribution in this paper is the most general of all the useful distributions applied in the BL-GARCH model parameter estimation. It offers a three parameter form which makes it more general than those available in the literature for obtaining empirical characteristics such as volatility clustering, leptokurtosis and leverage effect between returns and conditional variances as well as capturing heavier and lighter tails in high frequency financial time series data. The results of the extreme departure of some data are very crucial to risk managers in planning and decision making processes. Thus, the empirical results of BL-GARCH model show that parameters can be evaluated from Gaussian and non-Gaussian distributions.

REFERENCES

- [1] Abass, O. (1980). A Simulation Model for Electricity Consumption in Nigeria. *Recherche Di Automatica* II (2), 148-168.
- [2] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* AC-19: 716-723.
- [3] Bollerslev, T., Engle R. F. and Nelson D. B. (1994). ARCH Models in *Handbook of Econometrics*, vol. IV, edited by Engle R. F. and McFadden D. L. Elsevier Science.
- [4] Box, G.E.P. and Jenkins, G. W. (1976). *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.
- [5] Box, G. E. P. and Jenkins, G. M. and Reinsel, G. C. (1994). *Time Series Analysis, Forecasting and Control*, Prentice-Hall, Englewood Cliffs.
- [6] Cai, J. (1994). A Market Model of Switching-Regime ARCH. *Journal of Business and Economic Statistics*, 12, 309-316.
- [7] Christensen, R. (1996). *Plane Answers to Complex Questions: The Theory of Linear Models*, Second Edition, Springer-Verlag, New York Inc., 331.
- [8] Dallah, H., Okafor, R. O. and Abass, O. (2004). A Model-Based Bootstrap Method for Heteroskedasticity Regression Models, *Journal of Scientific Research and Development*, 9, 9 - 22.
- [9] Diongue, A. K., Guegan, D. and Wolff, R. C. (2010). Exact Maximum likelihood Estimator for the BL-GARCH model under elliptical distributed innovations. *Journal of Statistical Computation and Simulation*, 7, 4-10.
- Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of United Kingdom Inflation. *Econometrica*, 50, 987-1007.
- [10] Engle, R. F. (1990). Discussion: Stock Market Volatility and the Crash of '87. *Review of Financial Studies*, 3, 103-106.
- [11] Engle, R. F., Ng, V.K. and Rothschild (1990). Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills. *Journal of Econometrics*, 45, 213-238.
- [12] Engle, R. F. and Susmel R. (1993). Common Volatility in International Equity Markets. *Journal of Business and Economic Statistics*, 11, 167-176.
- [13] Eraker, B., Johannes, M. and Polson, N. (2003). The Impact of Jumps in Volatility and Returns. *Journal of Finance*. 58, 1269 & 1300.
- [14] Geweke, J. (1986). Modelling the persistence of Conditional Variances: Comment. *Econometric Reviews*, 5, 57-61.
- [15] Geweke, J. (1989). Exact Predictive Density for Linear Models with ARCH disturbances. *Journal of Econometrics*, 40, 63-86.
- [16] Gilks, W. and Berzuini, C. (2001), Following a Moving Target - Monte Carlo Inference for Dynamic Bayesian Model. *Journal of the Royal Statistical Society*, B63, 127 & 146.
- [17] Gouriéroux, C. (2007). *ARCH Models and Financial Applications*, Springer Series in Statistics, New York, 67-72.
- [18] Hamao, Y., Masulis, W. and Ng, V. K. (1990). Correlations in Price Changes and Volatility across International Stock Markets. *Review of Financial Studies*. 3, 281-307.
- [19] Jones, C. G., Kaul, G. K. and Lipson, M. L. (1994). Transactions, Volumes and Volatility. *Review of Financial Studies*, 7, 631-651.
- [20] Klein, B. (1977). The demand for quality-adjusted cash balances: Price uncertainty in the US demand for money functions. *Journal of Political Economy*, 85, 692-715.
- [21] Lamoureux, G. C. and Lastrapes, W. O. (1990). Heteroskedasticity in Stock Return Data: Volume Versus GARCH effects. *Journal of Finance*, 45, 221-229.
- [22] Linton, O. (1993). Adaptive Estimation in ARCH Models. *Econometric Theory*, 9, 539-569.
- [23] Lye, J. N., Martin, V. L. and Lesilie, E. T. (1998). Parametric Distributional Flexibility and Conditional Variance Methods with an Application to Hourly Exchange Rates. *International Monetary Fund (IMF)*, Policy Development and Review Department, WP/98/29, 1-39.
- [24] Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, 36, 394-419.

- [25] Mizrach, B. (1990). Learning and Conditional Heteroskedasticity in Asset Returns. Department of Finance, The Warthon School, University of Pennsylvania.
- [26] Mohler, R. R. (1973). Bilinear control processes – with applications to engineering, ecology and medicine. Academic Press, New York.
- [27] Onyeka-Ubaka, J. N., Abass, O. and Okafor, R. O. (2010). National Economic Growth, Poverty Eradication and the Statistician in Nigeria. *Nigeria Statistical Association (NSA) 2010 Conference Proceedings*, Imo Concorde Hotel, Owerri, Imo State, Pp. 171-183.
- [28] Onyeka-Ubaka, J. N. and Abass, O. (2013). Central Bank of Nigeria (CBN) Intervention and the Future of Stocks in the Banking Sector. *American Journal of Mathematics and Statistics*, 3(6): 407-416.
- [29] Pagan, A. R. and Schwert, G. W. (1990). Alternative Models for Conditional Stock Volatility, *Journal of Economics*, 45, 267-290.
- [30] Posedel P. (2005). Properties and Estimation of GARCH (1, 1) Model. *Metodoloski zvez* 2(2), 243-257.
- [31] Schwert, G. W. (1989). Why Does Market Volatility Change Over Time? *Journal of Finance*, 44, 1115-1153.
- [32] Storti, G., Vitale, C. (2003). BL-GARCH models and asymmetries in volatility. *Statistical Methods and Applications*, 12, 19-40.