

Conditional Variance Parameters in Symmetric 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 The parameters α and β are restricted to be non-negative in GARCH model, which have some consequences for the stationarity condition, and although the disturbances ε_t have mean 0, they are clearly not white noise because of their time-varying asymmetric probability density functions. Thus, this stationary process is capable of capturing well known phenomena present in financial markets such as volatility clustering, marginal distributions having heavy tails and thin centres (Leptokurtosis); return series appearing to be almost uncorrelated over time but to be dependent through higher moments. The possibility of having dependence between higher conditional moments, most notably variances, involves examining nonlinear stochastic processes from a more realistic perspective in time series data. This motivates the consideration of nonlinear models. The results obtained through Monte Carlo simulations established the practicability and small sample performance of symmetric models under Gaussian distributions. The results also showed that the corresponding standard errors are very small indicating that estimators are asymptotically unbiased, efficient and consistent at least within the sample.

Keywords GARCH, Symmetric, Simulation, Conditional Variance, Stationarity

1. Introduction

Financial time series present various forms of non linear dynamics, the crucial one being the strong dependence of the variability of the series on its own past and furthermore, with the fitting of the standard linear models being poor in these series. The autoregressive conditional heteroskedasticity (ARCH) model has become one of the most important models in financial applications. These models are non-constant variances conditioned on the past, which are a linear function on recent past disturbances Onyeka-Ubaka and Abass[9] and Dallah, Okafor and Abass[5]. This means that the more recent news will be the fundamental information that is relevant for modelling the present volatility. Moreover, the accuracy of the forecast over time improves when some additional information from the past is considered. Specifically, the conditional variance of the innovations will be used to calculate the percentile of the forecasting intervals, instead of the homoskedastic formula used in standard time series models.

Recent developments in economic and financial econometrics suggest the use of nonlinear time series structures to model the attitude of investors toward risk and expected returns. This is because the conditional variance is not constant over time. Hence, this paper, therefore, is set out

to establish the practicability and small sample performance of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model within the Gaussian framework.

2. Literature Review

Empirical studies show that models which present some nonlinearity can be modelled by conditional specifications, in both conditional mean and variance. A stochastic process

$\{y_t\}_{t=-\infty}^{\infty}$ is a model that describes the probability structure of a sequence of observations over time Zhu[13]. A time series y_t is a sample realization of a stochastic process that is observed only for a finite number of periods, indexed by $t = 1, 2, \dots, T$ Onyeka-Ubaka[8]. Any stochastic process can be partially characterized by the first and second moments of the joint probability distribution: the set of means, $\mu_t = E(y_t)$ and the set of variances and covariances $\text{cov}(y_t, y_h) = E(y_t - \mu_t)(y_h - \mu_h)$, $\forall t, h$. In order to get consistent forecast methods, we need that the underlying probabilistic structure would be stable over time. So a stochastic process is called weak stationary or covariance stationary when the mean, the variance and the covariance structure of the process is stable over time, that is:

$$E(y_t) = \mu < \infty$$

$$E(y_t - \mu)^2 = \sigma^2 < \infty$$

* Corresponding author:

jonyeka-ubaka@unilag.edu.ng (Onyeka-Ubaka J. N.)

Published online at <http://journal.sapub.org/ijps>

Copyright © 2014 Scientific & Academic Publishing. All Rights Reserved

Let $\{y_t(\theta)\}$ refer to the univariate discrete time-valued stochastic process to be predicted (e.g. the rate of return of a particular stock or market portfolio from time $t-1$ to t) where θ is a vector of unknown parameters and $E(y_t(\theta)|\psi_{t-1}) \equiv E_{t-1}(y_t(\theta)) \equiv \mu_t(\theta)$ denotes the conditional mean given the information set ψ_{t-1} available in time $t-1$. The innovation process for the conditional mean, $\{\varepsilon_t(\theta)\}$, is then given by $\varepsilon_t(\theta) = y_t(\theta) - \mu_t(\theta)$ with corresponding unconditional variance $V(\varepsilon_t(\theta)) = E(\varepsilon_t^2(\theta)) \equiv \sigma^2(\theta)$, zero unconditional mean and $E(\varepsilon_t(\theta)\varepsilon_h(\theta)) = 0$, $\forall t \neq h$. The conditional variance of the process given by ψ_{t-1} is defined by

$$V(y_t(\theta)|\psi_{t-1}) \equiv V_{t-1}(y_t(\theta)) \equiv E_{t-1}(\varepsilon_t^2(\theta)) \equiv \sigma_t^2(\theta).$$

Since investors would know the information set ψ_{t-1} when they make their investment decisions at time $t-1$, the relevant expected return to the investors and volatility are $\mu_t(\theta)$ and $\sigma_t^2(\theta)$, respectively.

An ARCH process, $\{\varepsilon_t(\theta)\}$, can be presented as:

$$\varepsilon_t(\theta) = z_t \sigma_t(\theta) \quad (1)$$

$$\stackrel{i.i.d}{z_t} \sim N(0, 1)$$

$$\begin{aligned} \sigma_t^2(\theta) &= g(\sigma_{t-1}(\theta), \sigma_{t-2}(\theta), \dots; \\ \varepsilon_{t-1}(\theta), \varepsilon_{t-2}(\theta), \dots; v_{t-1}, v_{t-2}, \dots) \end{aligned} \quad (2)$$

where $E(z_t) = 0$, $V(z_t) = 1$, $\sigma_t(\theta)$ is a time-varying positive and measurable function of the information set at time $t-1$, v_t is a vector of predetermined variables included in ψ_t , $g(\cdot)$ is a linear or nonlinear functional form. By definition, $\varepsilon_t(\theta)$ is serially uncorrelated with mean zero, but a time varying conditional variance equal to $\sigma_t^2(\theta)$. The conditional variance is a linear or nonlinear function of lagged values of σ_t and ε_t , and predetermined variables $(v_{t-1}, v_{t-2}, \dots)$ included in ψ_{t-1} . In the sequel, for notational convenience, no explicit indication of the dependence on the vector of parameters, θ , is given when obvious from the context.

Since very few financial time series have a constant conditional mean zero, an ARCH model can be presented in a regression form by letting ε_t be the innovation process in a linear regression:

$$y_t = x_t' b + \varepsilon_t \quad (3)$$

$$\varepsilon_t | \psi_{t-1} \sim N(0, \sigma_t^2)$$

where x_t is a $k \times 1$ vector of endogenous explanatory variables included in the information set ψ_{t-1} , b is a $k \times 1$ vector of unknown parameters, ε_t is the random disturbance term, and the subscript t indicates that X_t and Y_t are series of equally spaced observations through time. The use of the time series regression model underlies a number of assumptions concerning the form of the model, the independent variable and the disturbance terms. Provided that these assumptions hold:

(i) Zero mean: $E[e_t] = 0$

(ii) Constant Variance: $E[e_t^2] = \sigma^2$

(iii) Non-autoregression: $E[e_t e_{t-h}] = 0$ ($h \neq 0$)

It is possible to estimate optimally the regression parameters and their variances with the following formulas:

$$\hat{b} = \frac{\sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^T (X_t - \bar{X})^2} \quad (4)$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{X}_t \quad (5)$$

$$Var(\hat{b}) = \frac{s_e^2}{\sum_{t=1}^T (X_t - \bar{X})^2} \quad (6)$$

$$\begin{aligned} Var(\hat{a}) &= s_e^2 \left[\frac{1}{T} + \frac{\bar{X}^2}{\sum_{t=1}^T (X_t - \bar{X})^2} \right] \quad (7) \\ s_e^2 &= \frac{\sum_{t=1}^T (Y_t - \hat{Y})^2}{T - 2} \end{aligned}$$

where

$$\hat{Y} = \hat{a} + \hat{b}X_t \quad (8)$$

The estimators are optimal in the sense that they are unbiased, efficient and consistent. Unbiased estimators are those in which the expected value of the estimator, say \hat{b} , is equal to the true value, b . An estimator is relatively efficient if it has a smaller variance than any other estimator of b , and it will be consistent if both its bias and variance approach zero as the sample size approaches infinity. Together these properties mean that an estimator will be centered around the true value as the sample size increases. Of particular interest

are assumptions (i) and (iii), which, together, imply that the covariance of any two disturbance terms (i. e., $\text{cov}[e_t, e_{t-h}]$) is equal to zero. This can be seen as follows:

$$\begin{aligned}\text{Cov}[e_t, e_{t-h}] &= E[e_t - E(e_t)]E[e_{t-h} - E(e_{t-h})] \\ &= E[e_t - 0]E[e_{t-h} - 0] \\ &= E[e_t e_{t-h}] \\ &= 0\end{aligned}$$

This means that one assumes that disturbances at one point in time are not correlated with any other disturbances. The basic indicator of whether the non-autoregression assumption is violated is whether there is a sample correlation between the various random disturbance terms. To obtain a visual indication of the nature of the correlation, it is helpful to construct a correlogram that provides a graphical representation of the estimated autocorrelation function with time lags and autocorrelation coefficients forming the axes.

The paper observes that the main problem with an ARCH model of Engle[6] in empirical application is that it requires a large number of lags to catch the nature of the volatility; this can be problematic as it is difficult to decide how many lags to include and produces a non-parsimonious model where the non-negativity constraint could be failed.

To facilitate the computational problems of ARCH model, Bollerslev[1] proposed a generalization of the ARCH (q) process to allow for past conditional variances in the current equation.

$$\begin{aligned}E[\varepsilon_t | \mathcal{F}_{t-1}] &= 0, \\ \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2\end{aligned}\quad (9)$$

where

$$\omega > 0, \quad \alpha_i \geq 0, \quad i=1, 2, \dots, q, \quad \beta_j \geq 0, \quad j=1, 2, \dots, p$$

These conditions on parameters ensure strong positivity of the conditional variance (9). If the study writes the equation (9) in terms of lag-operator B, it gets

$$\sigma_t^2 = \omega + \alpha(B)\varepsilon_t^2 + \beta(B)\sigma_t^2 \quad (10)$$

where

$$\alpha(B) = \alpha_1 B + \alpha_2 B^2 + \dots + \alpha_q B^q$$

and

$$\beta(B) = \beta_1 B + \beta_2 B^2 + \dots + \beta_p B^p \quad (11)$$

The model is covariance stationary if all the roots of $\alpha(B) + \beta(B) = 1$ lie outside the unit circle, or

equivalently if $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$. That is, if

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

$$\sigma_t^2 - \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 - \sum_{j=1}^p \beta_j \sigma_{t-j}^2 = \omega$$

$$\sigma_t^2 \left(1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j \right) = \omega$$

Then, its long-run average variance (unconditional variance) is equal to

$$\sigma^2 \equiv \varepsilon^2 \equiv \frac{\omega}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j} \quad (12)$$

This model differs to the ARCH model in that it incorporates squared conditional variance terms as additional explanatory variables. This allows the conditional variance to follow an ARMA process in the squared innovations of orders $\max(p, q)$ and p , [ARMA($\max(p, q)$, p)], respectively:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \varepsilon_{t-j}^2 - \sum_{j=1}^p \beta_j v_{t-j} + v_t \quad (13)$$

The GARCH model is usually much more parsimonious and often a GARCH (1, 1) model is sufficient, this is because the GARCH model incorporates much of the information that a much larger ARCH model with large number of lags would contain.

3. Methodology

In this paper, the modelling is proceeded by specifying and estimating a model and then checking its adequacy. If model defects are detected at the latter stage, model revisions are made until a satisfactory model has been found. Then the model may be used for forecasting. Figure 1 depicts the main steps of a GARCH analysis and it is on these steps that this research is organized accordingly.

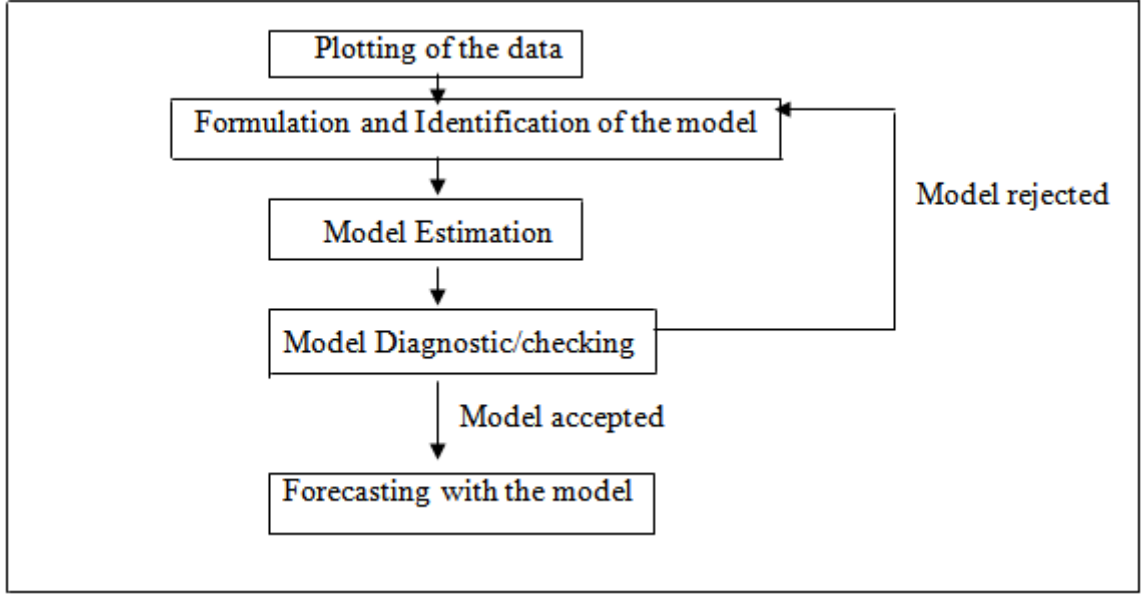


Figure 1. GARCH analysis (figure adapted from Box-Jenkins (1976) step by step approach)

4. Results and Discussion

The estimation of GARCH model involves the maximization of a likelihood function constructed under the auxiliary assumption of an independent identically distributed (i.i.d.) distribution for the standardized innovation $z_t(\theta)$. Let $f(z_t(\theta); \eta)$ denote the density function for $z_t(\theta) = \varepsilon_t(\theta) / \sigma_t(\theta)$, with mean zero and variance one, where η is the nuisance parameter, $\eta \in H \subseteq \Re^k$ is the vector of the parameter of f to be estimated.

To implement the maximum likelihood procedure, we use normal, the most commonly distribution in the literature.

$$f[z_t(\theta); \eta] = (2\pi)^{-1/2} \exp\left\{-\frac{z_t(\theta)^2}{2}\right\} \quad (14)$$

Since the normal distribution is uniquely determined by its first two moments, only the conditional mean and variance parameters enter the log-likelihood function in (9) i.e. $\psi = \theta$. The log-likelihood is

$$l_t = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} z_t(\theta)^2 - \frac{1}{2} \ln(\sigma_t^2)$$

It follows that the score vector $\left(s_t \equiv \frac{\partial l_t(y_t; \psi)}{\partial \psi}\right)$ takes the form:

$$\begin{aligned}
 s_t &= -z_t \frac{\partial z_t}{\partial \theta} - \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial(\sigma_t^2(\theta))}{\partial \theta} \\
 &= -\frac{\varepsilon_t(\theta)}{\sqrt{\sigma_t^2}} \frac{\partial(\varepsilon_t(\theta)(\sigma_t^2)^{-1/2})}{\partial \theta} - \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial(\sigma_t^2(\theta))}{\partial \theta} \\
 &= -\frac{\varepsilon_t(\theta)}{\sqrt{\sigma_t^2}} \frac{\partial((y_t - \mu_t(\theta))(\sigma_t^2)^{-1/2})}{\partial \theta} - \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial(\sigma_t^2(\theta))}{\partial \theta} \\
 &= \frac{\varepsilon_t(\theta)}{\sqrt{\sigma_t^2}} \frac{\partial \mu_t(\theta)}{\partial \theta} \sigma_t^2(\theta)^{1/2} + \frac{1}{2} (\sigma_t^2(\theta))^{-3/2} \frac{\partial(\sigma_t^2(\theta))}{\partial \theta} \frac{\varepsilon_t^2(\theta)}{\sqrt{\sigma_t^2(\theta)}} - \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial(\sigma_t^2(\theta))}{\partial \theta}
 \end{aligned}$$

$$\begin{aligned}
&= \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} + \frac{1}{2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\varepsilon_t^2(\theta)}{(\sigma_t^2(\theta))^2} - \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial(\sigma_t^2(\theta))}{\partial \theta} \\
&= \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} + \frac{1}{2} (\sigma_t^2(\theta))^{-1} \frac{\partial(\sigma_t^2(\theta))}{\partial \theta} \left[\frac{\varepsilon_t^2(\theta)}{\sigma_t^2(\theta)} - 1 \right]
\end{aligned} \quad (15)$$

When $\theta = (\alpha', \beta')$ where α are the conditional mean parameters and β are the conditional variance parameters, the score takes the form:

$$s_t = \begin{pmatrix} \frac{\partial l_t}{\partial \alpha} \\ \frac{\partial l_t}{\partial \beta} \end{pmatrix}$$

$$\text{where } \frac{\partial l_t}{\partial \alpha} = \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\varepsilon_t(\alpha)}{\sigma_t^2(\beta)}, \quad \frac{\partial l_t}{\partial \beta} = \frac{1}{2} (\sigma_t^2(\beta))^{-1} \frac{\partial \sigma_t^2(\beta)}{\partial \beta} \left[\frac{\varepsilon_t^2(\alpha)}{\sigma_t^2(\beta)} - 1 \right].$$

Even in the case of the symmetric GARCH (p, q) with normally distributed innovations, we have to solve a set of $\tilde{\theta} = k + p + q + 1$ nonlinear equations in (9). Numerical techniques are used in order to estimate the vector of parameters ψ . The problem faced in nonlinear estimation, as in the case of the GARCH models, is that there are no closed form solutions. So, an iterative method has to be applied to obtain a solution. Iterative optimization algorithms work by taking an initial set of values of the parameters, say $\psi^{(0)}$, then performing calculations based on these values to obtain a better set of parameters values $\psi^{(1)}$. The process is repeated until the likelihood function

$$S_T(y_T, y_{T-1}, \dots, y_1; \psi) \equiv \sum_{t=1}^T s_t(y_t; \psi) \quad (16)$$

no longer improves between iterations. If $\psi^{(0)}$ is a trial value of the estimate, then expanding $L_T(\{y_t\}; \psi) / \partial \psi$ and retaining only the first power of $\psi - \psi^{(0)}$, we obtain

$$\frac{\partial L_T}{\partial \psi} \approx \frac{\partial L_T}{\partial \psi^{(0)}} + (\psi - \psi^{(0)}) \frac{\partial^2 L_T}{\partial \psi^{(0)} \partial \psi'^{(0)}}.$$

At the maximum, $L_T / \partial \psi$ should equal zero. Rearranging terms, the correction for the initial value, $\psi^{(0)}$, obtained is

$$(\psi - \psi^{(0)}) = - \frac{\partial L_T}{\partial \psi^{(0)}} \left(\frac{\partial^2 L_T}{\partial \psi^{(0)} \partial \psi'^{(0)}} \right)^{-1}. \quad (17)$$

Let $\psi^{(i)}$ denote the parameter estimates after the i^{th}

iteration. Based on (17) the Newton-Raphson algorithm computes $\psi^{(i+1)}$ as:

$$\psi^{(i+1)} = \psi^{(i)} - \left(\frac{\partial^2 L_T^{(i)}}{\partial \psi \partial \psi'} \right)^{-1} \frac{\partial L_T^{(i)}}{\partial \psi}. \quad (18)$$

The scoring algorithm is a method closely related to the Newton-Raphson algorithm and was applied by Engle[6] to estimate the parameters of the ARCH (q) model. The difference between the Newton-Raphson method and the method of scoring is that the former depends on observed second derivatives, while the latter depends on the expected values of the second derivatives. So, the scoring algorithm computes $\psi^{(i+1)}$ as:

$$\psi^{(i+1)} = \psi^{(i)} + E \left(\frac{\partial^2 L_T^{(i)}}{\partial \psi \partial \psi'} \right)^{-1} \frac{\partial L_T^{(i)}}{\partial \psi}. \quad (19)$$

The assumption of normally distributed standardized innovations is often violated by the data. If the true distribution is instead leptokurtic, then the maximum is still consistent, but no longer efficient. In this case the maximum likelihood method is interpreted as the 'Quasi-Maximum Likelihood (QML) method. Bollerslev and Wooldridge[3], based on Wesis[12] and Pagan and Sabau[10], showed that the maximization of the normal log-likelihood function can provide consistent estimates of the parameter vector θ even when the distribution of z_t is non-normal, provided that

$$\begin{aligned}
E(z_t | \psi_{t-1}) &= 0 \\
E(z_t^2 | \psi_{t-1}) &= 1.
\end{aligned}$$

These estimates are, however, inefficient with the degree

of inefficiency increasing with the degree of departure from normality Pagan and Schwert[11]. So, the standard errors of the parameters have to be adjusted. Let $\hat{\theta}$ be the estimate that maximizes the normal log-likelihood function, in equation (16), based on the normal density function in (15) and let θ_0 be the true value. Then, even when z_t is non-normal, under certain regularity conditions:

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, A^{-1}BA^{-1})$$

Thus, $\hat{\theta}$ is a maximum likelihood estimator based on a mis-specified model. Under regularity conditions, the QML estimator converges almost surely to the pseudo true value θ_0 .

For symmetric departures from normality, the quasi-maximum likelihood estimation is generally close to the exact maximum likelihood estimation (MLE). But, for non-symmetric distribution, Engle and Kroner[7] showed that the loss in efficiency may be quite high. The practical applicability and small sample performance of the MLE procedure for the GARCH process is studied by Monte Carlo simulations.

4.1. Monte Carlo Experiments

A hybrid Monte Carlo experiment was performed using the normal distribution as data generating processes by introducing an auxiliary vector to avoid random walk behaviour. The momentum samples are discarded after sampling. The end result of hybrid Monte Carlo experiment is that the proposals move across the sample space in large steps and are therefore less correlated and converge to the target distribution more rapidly. Through the Monte Carlo

experiment, the model considered for $\varepsilon_t = y_t - \mu_t$ is a GARCH (1, 1) given by

$$Y_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = \sigma_t z_t \quad \text{for } t = 1, \dots, n,$$

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

where ε_t is a standard Normal random variable and $n = 50, 60, 70, 80, 250, 700, 1000$ and 3000 . The conditional mean, μ_t , is assumed to follow an AR (1) model in the simulation experiment.

Table 1 lists the Monte Carlo mean, standard error (Std error) and t-statistic for the parameter vector ω across $M = 1000$ Monte Carlo simulations. The simulation algorithm generates $n + 1000$ observations for each series, saving only the last n . This operation is performed in order to avoid dependence on initial values. The calculations were carried out in MATLAB (R2008b). For any set of parameter ω, α, β , the starting estimate for the variance of the first observation is often taken to be the observed variance of the residuals. It is easy to calculate the variance forecast for the second. The GARCH updating formula takes the weighted average of the unconditional variance, the squared residuals for the first observation and the starting variance and estimates of the second. This is input into the forecast of the third variance, and so forth. Ideally, this series is large when the residuals are large and small when they are small. The likelihood function provides a systematic way to adjust the parameters ω, α, β to give the best fit.

Table 1. Estimated Parameters (and the Estimated Standard Errors) for the Centered Normal GARCH (1, 1) Model

N	$\hat{\alpha}_0$			$\hat{\alpha}_1$			$\hat{\beta}_1$		
	Mean	Std error	t-statistic	Mean	Std error	t-statistic	Mean	Std error	t-statistic
50	-0.0450 (0.4218)	0.7237	-0.0622	0.2301 (0.0376)	0.3681	0.6251	1.0000 (0.0100)	0.0896	11.1607
60	-1.0535 (0.5901)	1.0653	-0.9890	1.0988 (0.0145)	0.4992	2.2011	0.8185 (0.0762)	0.1886	4.3399
70	-0.4197 (0.0190)	2.8838	-0.1455	0.2709 (0.0843)	0.5175	0.5234	0.2363 (0.2715)	1.0779	0.2192
80	-0.0051 (0.0245)	0.5415	-0.0946	-0.2048 (0.0265)	0.3026	-0.6768	0.3571 (0.0569)	0.2780	1.2845
250	-0.0067 (0.0147)	0.0768	-0.0872	0.1642 (0.0039)	0.0752	2.1835	0.9447 (0.0025)	0.0499	18.9319
700	-0.0046 (0.0284)	0.0590	-0.0780	0.2151 (0.0247)	0.0555	3.8757	0.9812 (0.0017)	0.0103	95.2621
1000	-0.0022 (0.0486)	0.0388	-0.0567	0.2999 (0.00879)	0.0401	7.4788	0.9809 (0.0009)	0.0076	129.0657
3000	-0.0016 (0.0293)	0.0296	-0.0541	0.2510 (0.0057)	0.0203	12.3645	0.9895 (0.0004)	0.0030	329.8333

Inspection of Table 1 reveals that for all sample sizes, the averages obtained from the exact MLE are close to the true parameter values. The corresponding standard errors are very small indicating that estimators are asymptotically unbiased, efficient and consistent. The standard GARCH model has news impact curve, which is symmetric and centred at $\varepsilon_{t-1} = 0$. That is, positive and negative return shocks of the same magnitude produce the same amount of volatility. Also, larger return shocks forecast more volatility at a rate proportional to the square of the size of the return shock. If a negative return shock causes more volatility than a positive return shock of the same size, the GARCH model under predicts the amount of volatility following bad news and over predicts the amount of volatility following good news. Furthermore, if large return shocks cause more volatility than a quadratic function allows, then the standard GARCH model under predicts volatility after a large return shock and over predicts volatility after a small return shock.

Table 1 summarizes the estimates coefficients from the GARCH (1, 1) model with the true value set of parameters $\{\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1, \hat{c}_1\} = \{-0.0016, 0.2510, 0.9895\}$, that is, the simulation model: $\sigma_t^2 = -0.0016 + 0.2510\varepsilon_{t-1}^2 + 0.9895\sigma_{t-1}^2$. The P-values of the estimated parameters shown in the brackets are statistically significant with values almost less than 0.05. Where Std error is the Standard errors and t-statistic is T Statistics obtained by dividing the mean by the standard error. Monte Carlo simulations are computed 1000 replications. Each replication gives a sample size $n = 50, 60, 70, 80, 250, 700, 1000$ and 3000 observations.

5. Conclusions

The GARCH model is characterized by a symmetric response of current volatility to positive and negative lagged errors ε_{t-1} , since ε_t is uncorrelated with its history. This could be interpreted conveniently as a measure of news entering a financial market at time t . The paper established that the GARCH (1, 1) model captures volatility clustering and leptokurtosis present in high frequency financial time series data. The Monte Carlo simulations showed the practicability and sample performance analysis of the GARCH (1, 1) model under normal distribution. The simulated results also showed that the corresponding standard errors are very small indicating that estimators are asymptotically unbiased, efficient and consistent. From the discussions, we observe that the simple structure of the model imposes important limitations on GARCH models. Thus:

(i) GARCH models, however, assume that only the magnitude and not the positivity or negativity of unanticipated excess returns determines conditional variance σ_t^2 . If the distribution of z_t is symmetric, the change in variance tomorrow is conditionally uncorrelated with excess returns today.

(ii) The GARCH models are not able to explain the

observed covariance between ε_t^2 and ε_{t-j} . This is possible only if the conditional variance is expressed as an asymmetric function of ε_{t-j} .

(iii) GARCH models essentially specify the behaviour of the square of the data. In this case a few large observations can dominate the sample.

(iv) The negative correlation between stock returns and changes in returns volatility, that is, volatility tends to rise in response to “bad news”, (excess returns lower than expected) and to fall in response to “good news” (excess returns higher than expected) are also observed. This limitation is called leverage effect and can be tackled with asymmetric models.

REFERENCES

- [1] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307-327.
- [2] Bollerslev, Engle and Wooldridge (1988). A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy*. 96, 116-131.
- [3] Bollerslev, T. and Wooldridge, J. M. (1992). Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances. *Econometric Reviews*, 11, 143-173.
- [4] Box, G.E.P. and Jenkins, G. W. (1976). Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco.
- [5] 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.
- [6] Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of United Kingdom Inflation. *Econometrica*, 50, 987-1007.
- [7] Engle, R. F. and Kroner, K. F. (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 11, 122-150.
- [8] Onyeka-Ubaka, J. N. (2013). A Modified BL-GARCH Model for Distributions with Heavy Tails. A Ph.D Thesis, University of Lagos, Akoka, Nigeria.
- [9] 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.
- [10] Pagan, A. R. and Sabau, H. (1987). On the Inconsistency of the MLE in Certain Heteroskedastic Regression Models, University of Rochester, Department of Economics, Mimeo.
- [11] Pagan, A. R. and Schwert, G. W. (1990). Alternative Models for Conditional Stock Volatility, *Journal of Economics*, 45, 267-290.
- [12] Weiss, A. (1986). Asymptotic Theory for ARCH Models: Estimation and Testing. *Econometric Theory*, 2, 107-131.
- [13] Zhu, F. (2011). A negative binomial integer-valued GARCH model. *Journal of Time Series Analysis*, 32, 54-67.