Estimation of Asymmetric Garch Models: The Estimating Functions Approach

Mr. Timothy Ndonye Mutunga
Prof. Ali Salim Islam
Dr. Luke Akong'o Orawo
Department of Mathematics
Egerton University
Egerton
Kenya

Abstract

This paper introduces the method of estimating functions (EF) in the estimation of the Asymmetric GARCH family of models. This approach utilises the third and fourth moments which are common in financial time series data analysis and does not rely on distributional assumptions of the data. Optimal estimating functions have been constructed as a combination of linear and quadratic estimating functions. Estimates from the estimating functions approach are better than those of the traditional estimation methods such as the maximum likelihood estimation (MLE) especially in cases where distributional assumptions on the data are highly violated. We investigate the presence of asymmetric (leverage) effects in empirical time series and fit two of the most popular Asymmetric GARCH models (EGARCH and GJR-GARCH) under both the MLE and EF approaches. An empirical example demonstrates the implementation of the EF approach to Asymmetric GARCH models assuming a student's – t distribution for the innovations. The efficiency benefits of the EF approach relative to the MLE method in parameter estimation are substantial for non-normal cases.

Keywords: Estimating function, Maximum likelihood estimation, Asymmetric GARCH, Volatility, Leverage effects

1. Introduction

The financial stock market has been an area of great interest to researchers for the last five decades. Starting with the pioneering works on the random walk model, stock market volatility has been a key subject in most subsequent research works relating to efficiency in the market (Fama, 1965; Fama, 1970). In financial decision making, volatility is an important factor in pricing of derivatives and portfolio risk management. This has warranted increased research on modelling and forecasting an asset's price/returns volatility.

Research on changing volatility using non-linear time series models has been vibrant since the introduction of the Autoregressive Conditional Heteroscedasticity (ARCH) model (Engle, 1982). This model was the first of its kind to takeconditional heteroscedasticity into consideration. Bollerslev (1986) generalised the ARCH model to include lagged conditional variances as well as lagged values of the squared innovations. The GARCH family of models have proved to be successful in capturing volatility clustering and some amount of the excess kurtosis which characterize financial time series data.

Since the works of Engle (1982) and Bollerslev (1986), various variants of the GARCH model have been developed to model volatility. Of great importance is the Asymmetric GARCH family of models which address a major limitation of the Bollerslev's (1986) basic GARCH model, relating to the inability of this model to capture the asymmetric impact of news on volatility. News is undoubtly a huge factor that affects stock prices and therefore measuring its impact on stock market volatility is an important area of research in financial theory (Neelabh, 2009). Different volatility models that capture this aspect have been proposed and widely applied to real life problems in the last two decades. Some of the most popular models include the EGARCH (Nelson, 1991), GJR-GARCH (Glosten*et al.*, 1993), NAGARCH (Engle and Ng, 1993), APARCH (Ding *et al.*, 1993), TGARCH (Zakoian, 1994) and the QGARCH (Sentana, 1995).

In the bulk of literature available for the Asymmetric GARCH models, the maximum likelihood estimation method has been the most preferred in parameter estimation due to its simplicity and desirable properties. However, this method is based on distributional assumptions which are often violated in practice and thus alternative parameter estimation approaches are required. An alternative method of estimation is based on the estimating functions (EF) approach introduced by Godambe (1960). Under this approach, focus is usually on the estimating function itself which is a function of the observations and the unknown parameters. This approach takes into account higher order moments and does not rely on any distributional assumptions on the data for optimality.

The focus of this paper is twofold. First we seek to introduce the method of estimating functions in the estimation of the Asymmetric GARCH family of models based on Godambe and Thompson's (1989) optimal estimating functions for stochastic processes. Secondly we will utilise the EF method in the estimation of first order EGARCH and GJR-GARCH models based on empirical time series from the USA and Japanese stock markets. An overview of these two Asymmetric GARCH models is presented in section 2. The computation of optimal estimating functions for the first order EGARCH and GJR-GARCH models and the Asymmetric GARCH – class of models in general, is presented in section 3. An example involving the two empirical financial time series is presented in section 4 to demonstrate the use of the EF approach in estimation of Asymmetric GARCH models particularly in cases where there are serious departures from normality. Finally a conclusion of this paper is presented in section 5.

2. Asymmetric GARCH – Class of Models

The conventional GARCH model besides its main virtue of simplicity imposes a number of shortcomings with regard to volatility modelling. However the primary limitation of the GARCH model relates to what Black (1976) first documented. There exists a negative correlation between stock returns and volatility implying that negative returns tend to be followed by larger increases in volatility while positive returns of the same magnitude tend to be followed by lower volatility. To model this phenomenon, this paper will consider two of the most popular models that allow for asymmetric shocks.

2.1 EGARCH Model

The EGARCH model was introduced by Nelson (1991) to address some of the weaknesses of the conventional GARCH model introduced by Bollerslev (1986). This model captures asymmetric responses of the conditional variance to shocks in the market. The variance equation in EGARCH (p, q) is specified as;

$$\ln h_t = \kappa + \sum_{i=1}^{p} \beta_i \ln h_{t-i} + \sum_{i=1}^{q} \alpha_i g\left(z_{t-i}\right) \tag{1}$$

where, $\alpha_1 = 1,\ \varepsilon_t = z_t\sqrt{h_t},\ \mathrm{g}\left(z_t\right) = \gamma_1 z_t + \gamma_2\left[\left|z_t\right| - E\left|z_t\right|\right]$.

The left hand side is the log of the variance series. This makes the leverage effect exponential and therefore the parameters $\kappa,\ \beta_i,$ and α_i are not restricted to be non-negative. α_i is the asymmetry parameter.

The value of $\mathrm{g}\left(z_t\right)$ must be a function of both magnitude and sign of z_t in order to accommodate the asymmetric effect (Nelson, 1991). The components $\gamma_1 z_t$ and $\gamma_2\left[\left|z_t\right| - E\left|z_t\right|\right]$ represent the sign effect and magnitude effect respectively and each has a zero mean.

Over the range $0 < z_t < \infty$, $\mathrm{g}\left(z_t\right)$ is linear in z_t with slope $\gamma_1 + \gamma_2$ while over the range $-\infty < z_t < 0$, $\mathrm{g}\left(z_t\right)$ is linear in z_t with slope $\gamma_1 - \gamma_2$. Thus $\mathrm{g}\left(z_t\right)$ allows the conditional variance h_t to respond asymmetrically to changes in stock returns. Consider the first order EGARCH model in (2),

$$\left.\begin{aligned} \varepsilon_t &= z_t\sqrt{h_t} \\ \ln h_t &= \kappa + \beta \ln h_{t-1} + \gamma_1 z_{t-1} + \gamma_2\left[\left|z_{t-1}\right| - E\left|z_{t-1}\right|\right] \end{aligned}\right\} \tag{2}$$

where, z_t is an identically distributed sequence of random variables with zero mean and a unit variance.

The term $g(z_{t-1}) = \gamma_1 z_{t-1} + \gamma_2\left[|z_{t-1}| - E|z_{t-1}|\right]$ gives the model capacity to capture asymmetry. If $\gamma_1 = 0$ and $\gamma_2 > 0$, the innovation (disturbance) in $\ln h_t$ is now positive (negative) when the magnitude of z_t is larger (smaller) than its expected value. On the other hand if $\gamma_1 < 0$ and $\gamma_2 = 0$ the innovation in $\ln h_t$ is now positive (negative) when returns innovations are negative (positive). Thus the EGARCH (1,1) model is able to capture the leverage effects under these conditions.

2.2 GJR-GARCH model

This model was introduced by Glosten *et al*., (1993). It is an extension of the GARCH model to capture asymmetries between positive and negative shocks of the same magnitude on the volatility of returns. The GJR-GARCH(p, q) model is specified as;

$$\left.\begin{aligned} \varepsilon_t &= z_t\sqrt{h_t} \\ h_t &= \kappa + \sum_{i=1}^{p}\left(\alpha_i\varepsilon_{t-i}^2 + \gamma s_{t-i}^{-}\varepsilon_{t-i}^2\right) + \sum_{i=1}^{q}\beta_i h_{t-i} \\ s_t^{-} &= \begin{cases} 1 & \text{when } \varepsilon_t < 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}\right\} \tag{3}$$

where, $z_t \sim iid\ N(0,1)$.

Consider the first order GJR-GARCH model in (4);

$$h_t = \kappa + \alpha\varepsilon_{t-1}^2 + \gamma s_{t-1}^{-}\varepsilon_{t-1}^2 + \beta h_{t-1} \tag{4}$$

The model reduces to the traditional GARCH model whenever $\varepsilon_t \geq 0$. The indicator term s_t^{-} ensures that the asymmetric effect is captured in the model. With $\gamma > 0,$ negative shocks $(\varepsilon_{t-1} < 0)$ increase volatility more than positive shocks $(\varepsilon_{t-1} > 0)$ of equal magnitude. The necessary and sufficient conditions to guarantee positivity of the conditional variance h_t are $\kappa > 0,\ \alpha > 0,\ \beta > 0$ and $\alpha + \gamma > 0.$ γ is the asymmetry parameter.

3. Optimal Estimating Functions

In this section we derive optimal estimating functions and show their application to Asymmetric GARCH family of models. We draw extensively on the works of Godambe (1960), Godambe (1985) and Godambe and Thompson (1989).Without proof we first present an important theorem related to the theory of EF's in stochastic processes.

Godambe and Thompson (1989) extended the concept of optimality of Godambe's (1985) EF into a general setting using a more flexible conditioning method which is related to the concept of quasi-likelihood approach. Taking Υ as an arbitrary sample space, they considered the class of EFs M_j which is a real function defined on $\Upsilon \times \Theta$ such that;

$$E\left[\mathrm{M}_j\left(y_1, y_2, ..., y_n; \theta\right) / \Upsilon_j\right] = 0 \tag{5}$$

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

Theorem: To estimate $\theta \in \Theta$ consider a class of EFs$\mathcal{H} = \{h\}$ for which;

$$h = \sum_{j=1}^{k} a_j \mathrm{M}_j \tag{6}$$

where, a_j is a real function on $\Upsilon \times \Theta$.

The EFs $\mathrm{M}_j,\ j=1,\ldots,k$ are said to be orthogonal if,

$$E\left(\mathrm{M}_j\mathrm{M}_i / \Upsilon_i\right) = E\left(\mathrm{M}_i\mathrm{M}_j / \Upsilon_j\right) = 0 \ \ for \ \ i \neq j, i, j = 1,\ldots,k. \tag{7}$$

An estimate $\hat{\theta}$ of θ is obtained by solving the equation $h\left(y_1, y_2, \ldots, y_n; \theta\right) = 0$. The optimal EF in this case is defined as;

$$h^* = \sum_{j=1}^{k} a_j^* \mathrm{M}_j \tag{8}$$

where,

$$a_j^* = \frac{E\left\{\left(\frac{\partial \mathrm{M}_j}{\partial \theta}\right) / \Upsilon_j\right\}}{E\left\{\left[\mathrm{M}_j\left(y_1, y_2, \ldots, y_n; \theta\right)\right]^2 / \Upsilon_j\right\}} \tag{9}$$

3.1 Parameter Estimation Using the Estimating Functions Approach

To estimate parameters of the EGARCH and GJR-GARCH models in a regression model set up using the EFs approach, optimal estimating functions approach to discrete time stochastic processes by Godambe and Thompson (1989) was applied.

Consider a general expression of the EGARCH and GJR-GARCH models in a regression model set up without making any distributional assumptions for the errors,

$$\left.\begin{aligned} y_t &= x_t\omega + \varepsilon_t \\ y_t / \psi_{t-1} &\sim \left(x_t\omega, h_t\right) \end{aligned}\right\} \tag{10}$$

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

Consider the first EGARCH model in (11).

$$\left.\begin{aligned} \varepsilon_t &= z_t\sqrt{h_t} \\ \ln h_t &= \kappa + \beta \ln h_{t-1} + \gamma_1 z_{t-1} + \gamma_2\left[\left|z_{t-1}\right| - E\left|z_{t-1}\right|\right] \\ h_t &= \exp\left\{\kappa + \beta \ln h_{t-1} + \gamma_1 z_{t-1} + \gamma_2\left[\left|z_{t-1}\right| - E\left|z_{t-1}\right|\right]\right\} \end{aligned}\right\} \tag{11}$$

where,

$\alpha_1 = 1$, $\mathrm{g}\left(z_t\right) = \gamma_1 z_t + \gamma_2\left[\left|z_t\right| - E\left|z_t\right|\right]$ and z_t is an independent and identically distributed sequence of random variables.

Let $\theta_1 = (\kappa, \beta, \gamma_1, \gamma_2)$. We seek to estimate the unknown parameter vectors ω and θ_1 in the regression model (10) where h_t is as defined in (11).

Similarly consider the first order GJR-GARCH model in (12),

$$\left.\begin{aligned} \varepsilon_t &= z_t\sqrt{h_t} \\ h_t &= \kappa + \alpha\varepsilon_{t-1}^2 + \gamma s_{t-1}^- \varepsilon_{t-1}^2 + \beta h_{t-1} \\ s_t^- &= \begin{cases} 1 & \text{when } \varepsilon_t < 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}\right\} \tag{12}$$

where, z_t is an independent and identically distributed sequence of random variables.

Let $\theta_2 = (\kappa, \alpha, \beta, \gamma)$. Similarly we seek to estimate the unknown parameter vectors ω and θ_2 in the regression model (10) where h_t is as defined in (12).

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

$$\left.\begin{aligned} \lambda_{1t} &= y_t - x_t\omega \\ \lambda_{2t}^* &= (y_t - x_t\omega)^2 - h_t \end{aligned}\right\} \tag{13}$$

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

$$E\left(\lambda_{1t}\ \lambda_{2t}^*\right) \neq 0 \tag{14}$$

λ_{2t}^* is therefore orthogonalised using the Gram –Schmidt orthognalisation procedure (Hyde, 1997) as follows,

$$\begin{aligned} \lambda_{2t}^* &= (y_t - x_t\omega)^2 - h_t \\ &= \left\{ E\left[(y_t - x_t\omega)^3 - h_t\,(y_t - x_t\omega)/\psi_{t-1}\right]\left[E(y_t - x_t\omega)^2/\psi_{t-1}\right]^{-1}(y_t - x_t\omega)/\psi_{t-1} \right\} \\ &= (y_t - x_t\omega)^2 - h_t - (y_t - x_t\omega)\,E\left\{\frac{(y_t - x_t\omega)^3}{h_t}/\psi_{t-1}\right\} \end{aligned} \tag{15}$$

From (15), consider the component (16),

$$E\left\{\frac{(y_t - x_t\omega)^3}{h_t}/\psi_{t-1}\right\} \tag{16}$$

Dividing and multiplying (16) by $\sqrt{h_t}$ we have,

$$E\left\{\frac{(y_t - x_t\omega)^3}{h_t^{3/2}}/\psi_{t-1}\right\} h_t^{1/2} = \phi_{1t}\ h_t^{1/2} \tag{17}$$

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

Thus,

$$\lambda_{2t} = (y_t - x_t\omega)^2 - h_t - \varepsilon_t\ \phi_{1t}\ h_t^{1/2} \tag{18}$$

Therefore the two elementary EFs in (19) are now orthogonal.

$$\left.\begin{aligned} \lambda_{1t} &= y_t - x_t\omega \\ \lambda_{2t} &= (y_t - x_t\omega)^2 - h_t - \varepsilon_t\ \phi_{1t}\ h_t^{1/2} \end{aligned}\right\} \tag{19}$$

To estimate the coefficient vectors ω and θ in the regression model (10), optimal EFs are derived using the elementary EFs in (19). The theorem by Godambe and Thompson (1989) is applied to form a linear combination of the elementary EFs as,

$$\left.\begin{aligned} g_1 &= \sum_{t=1}^{T} a_{1t}\lambda_{1t} + \sum_{t=1}^{T} a_{2t}\lambda_{2t} \\ g_2 &= \sum_{t=1}^{T} b_{1t}\lambda_{1t} + \sum_{t=1}^{T} b_{2t}\lambda_{2t} \end{aligned}\right\} \tag{20}$$

Let $\mathcal{L}$ be the class of all EFs $\left(g_{1,}g_2\right)$ given by (20). The jointly optimal EFs $\left(g^*_{1,}g^*_2\right)$ are given by (20) with, $a_{it}=a^*_{it}$ and $b_{it}=b^*_{it}$ for $i=1,2$ and $t=1,2,3,...,T$

Where,

$$b^*_{1t}=\frac{E\left(\frac{\partial\lambda_{1t}}{\partial\omega}/\psi_{t-1}\right)}{E\left(\lambda^2_{1t}/\psi_{t-1}\right)}=\frac{\frac{\partial x_t\omega}{\partial\omega}}{h_t} \tag{21}$$

$$b^*_{2t}=\frac{E\left(\frac{\partial\lambda_{2t}}{\partial\omega}/\psi_{t-1}\right)}{E\left(\lambda^2_{2t}/\psi_{t-1}\right)}$$

$$=\frac{E\frac{\partial}{\partial\omega}\left[\left(y_t-x_t\omega\right)^2-h_t-\left(y_t-x_t\omega\right)\phi_{1t}h_t^{1/2}/\psi_{t-1}\right]}{E\left(\lambda^2_{2t}/\psi_{t-1}\right)} \tag{22}$$

Solving the numerator in (22) we have,

$$E\frac{\partial}{\partial\omega}\left[\left(y_t-x_t\omega\right)^2-h_t-\left(y_t-x_t\omega\right)\phi_{1t}h_t^{1/2}/\psi_{t-1}\right]$$

$$=E\left\{-2\left(y_t-x_t\omega\right)\frac{\partial x_t\omega}{\partial\omega}-\frac{\partial h_t}{\partial\omega}-\frac{\left(y_t-x_t\omega\right)}{2h_t^{1/2}}\frac{\partial h_t}{\partial\omega}\phi_{1t}+\phi_{1t}h_t^{1/2}\frac{\partial x_t\omega}{\partial\omega}\right\}$$

$$=\phi_{1t}h_t^{1/2}\frac{\partial x_t\omega}{\partial\omega}-\frac{\partial h_t}{\partial\omega} \tag{23}$$

Solving the denominator in (22) we have,

$$E\left(\lambda^2_{2t}/\psi_{t-1}\right)$$

$$\begin{aligned}=E\Big\{\Big[&\left(y_t-x_t\omega\right)^4-2h_t\left(y_t-x_t\omega\right)^2+h_t^2+\phi_{1t}h_t^{3/2}\left(y_t-x_t\omega\right)-\phi_{1t}h_t^{1/2}\left(y_t-x_t\omega\right)^3\\ &+\phi_{1t}h_t^{3/2}\left(y_t-x_t\omega\right)\ -\phi_{1t}h_t^{1/2}\left(y_t-x_t\omega\right)^3+\phi^2_{1t}h_t\left(y_t-x_t\omega\right)^2\Big]/\psi_{t-1}\Big\}\end{aligned} \tag{24}$$

Multiplying and dividing (24) by h_t^2 leads to,

$$\begin{aligned}=Eh_t^2\Bigg\{\Bigg[&\frac{\left(y_t-x_t\omega\right)^4}{h_t^2}-\frac{2h_t\left(y_t-x_t\omega\right)^2}{h_t^2}+1+\frac{\phi_{1t}\left(y_t-x_t\omega\right)}{h_t^{1/2}}-\frac{\phi_{1t}\left(y_t-x_t\omega\right)^3}{h_t^{3/2}}+\frac{\phi_{1t}\left(y_t-x_t\omega\right)}{h^{1/2}}\\ &-\frac{\phi_{1t}\left(y_t-x_t\omega\right)^3}{h_t^{3/2}}+\frac{\phi^2_{1t}\left(y_t-x_t\omega\right)^2}{h_t}\Bigg]/\psi_{t-1}\Bigg\}\end{aligned}$$

$$=h_t^2\left(\phi_{2t}+2-\phi^2_{1t}\right) \tag{25}$$

Where,

$$\phi_{2t} = E\left\{\left[\frac{\left(y_t - x_t\omega\right)^4}{h_t^2} - 3\right] / \psi_{t-1}\right\} \tag{26}$$

Equation (26) represents the standardized conditional kurtosis (excess kurtosis).

Hence,

$$b_{2t}^* = \frac{\phi_{1t} h_t^{1/2} \dfrac{\partial x_t\omega}{\partial\omega} - \dfrac{\partial h_t}{\partial\omega}}{h_t^2\left(\phi_{2t} + 2 - \phi_{1t}^2\right)} \tag{27}$$

$$a_{1t}^* = \frac{E\left(\dfrac{\partial\lambda_{1t}}{\partial\theta} / \psi_{t-1}\right)}{E\left(\lambda_{1t}^2 / \psi_{t-1}\right)} = 0 \tag{28}$$

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

$$= \frac{E\dfrac{\partial}{\partial\theta}\left[\left(y_t - x_t\omega\right)^2 - h_t - \left(y_t - x_t\omega\right)\ \phi_{1t}\ h_t^{1/2}\right] / \psi_{t-1}}{E\left(\lambda_{2t}^2 / \psi_{t-1}\right)}$$

$$= \frac{E\left[-\dfrac{\partial h_t}{\partial\theta} - \dfrac{1}{2}\phi_{1t}\dfrac{\left(y_t - x_t\omega\right)}{h^{1/2}}\dfrac{\partial h_t}{\partial\theta}\right] / \psi_{t-1}}{E\left(\lambda_{2t}^2 / \psi_{t-1}\right)}$$

$$= \frac{-\dfrac{\partial h_t}{\partial\theta}}{h_t^2\left(\phi_{2t} + 2 - \phi_{1t}\right)} \tag{29}$$

Substituting (21), (27), (28) and (29) into (20) gives the jointly optimal EFs as,

$$\left.\begin{aligned} g_1^* &= -\sum_{t=1}^{T}\frac{\dfrac{\partial h_t}{\partial\theta}}{h_t^2\left(\phi_{2t} + 2 - \phi_{1t}\right)}\lambda_{2t} \\ g_2^* &= \sum_{t=1}^{T}\frac{\dfrac{\partial x_t\omega}{\partial\omega}}{h_t}\lambda_{1t} - \sum_{t=1}^{T}\frac{\dfrac{\partial h_t}{\partial\theta}}{h_t^2\left(\phi_{2t} + 2 - \phi_{1t}\right)}\lambda_{2t} \end{aligned}\right\} \tag{30}$$

Where, h_t is given by equations (11) and (12) and

$$\theta = \begin{cases} \theta_1 & for\ EGARCH\,(1,1) \\ \theta_2 & for\ GJR - GARCH\,(1,1) \end{cases} \tag{31}$$

The result in (30) is very general in that no distributional assumptions on y_t / ψ_{t-1} have been made.

The estimates for the unknown parameter vectors ω and θ are obtained by solving the optimal EF in (32). This means numerically minimizing $g_1^* + g_2^*$.

$$g^*_{\theta,\omega} = g^*_1 + g^*_2 = 0 \tag{32}$$

Where, g^*_1 and g^*_2 are as defined in (30).

4. Estimation of Asymmetric GARCH Models on Empirical Time Series

This section presents estimation results. A brief description of the real data that was used to fit the models is provided. Some preliminary and diagnostic tests for asymmetry and normality are also conducted beforehand.

4.1 Data

In this sub-section we model the volatility of financial returns of the Japanese and the USA markets for the period 2nd Jan 2008 to 31stMay 2011 using Asymmetric GARCH models under both the MLE and EF procedures. In each market we consider a comprehensive and diversified stock index. For the New York Stock Exchange we consider the Standard and Poor's 500 index while for the Tokyo Stock Exchange we consider the Nikkei 225 index. In each case daily returns are computed as logarithmic price (P_t) relatives.

$$R_t = \log \frac{P_t}{P_{t-1}} \tag{33}$$

where, $\{R_t\}$ is the log return series (continuously compounded return).

Each empirical time series comprises of daily observations covering the period (2nd Jan 2008 to 31st May 2011).Stock markets in these two countries were among the most volatile in the world during the 2008 global financial crises and the early 2011 Japanese tsunami disaster respectively and hence the impact of shocks in the market on volatility of asset returns was more pronounced during this period.

4.2 Preliminary Tests

Table 1 presents summary statistics (empirical properties) and preliminary tests of normality and asymmetry for the daily stock returns of the two financial series. We notice that the daily volatility for the Nikkei 225 index, represented by the standard deviation (1.88%) is above the volatility (1.75%) for the S&P 500 index return series.

Table 1: Summary Statistics of the compounded returns R_t

Series Statistics	**SP500 INDEX**	**NIKKEI 225 INDEX**
Mean	-0.000060	-0.000917
Std Dev.	0.017501	0.018802
Skewness	-0.163770	-1.022632
Kurtosis	7.422417	8.238645
Jarque – Bera (Probability) h	0.0000 1	0.0000 1
ADF test (Probability) h	1.000E-3 1	1.000E-3 1
$\text{Cov}\left(r_t^2, r_{t-1}\right)$	-0.078035	-0.064877

The skewness coefficient is negative for both series suggesting that they a long left tail. The kurtosis coefficient on the other hand is very high (7.4224, 8.2386) for the (S&P 500, Nikkei 225) a reflection that the distributions of the two sets of real data are highly leptokurtic. The P-value corresponding to the Jarque –Bera normality test is zero at 5% level suggesting that the test is significant for both series. The test returns a value of $h=1$ which indicates that the series R_t does not come from a normal distribution in favour of $h=0$ which indicates that the series R_t comes from a normal distribution with unknown mean and variance. This implies that the two series exhibit non-normal behaviour.

The Augmented Dickey-Fuller (ADF) test rejects the unit root null in both data sets. This is indicated by the minimal p-values at 5% level and the values of h. The test returns a value of $h = 1$ which indicates rejection of the unit root in favour of the trend-stationary alternative. $h = 0$ indicates failure to reject the unit root null. Thus we conclude that the returns of both stock indices are stationary.

Finally we test for presence of asymmetric effects on conditional volatility in both empirical series. A simple diagnostic test for the leverage effects involves computing the sample correlation between squared returns and the lagged returns, $\left\{\mathrm{Cov}\left(r_t^2, r_{t-1}\right)\right\}$ (Zivot, 2008). A negative value for this coefficient provides evidence for potential asymmetric effects. Both series have a negative value for this coefficient indicating evidence of asymmetry and hence Asymmetric GARCH family of models could perform well in explaining conditional volatility in this case.

Figure 1: Daily Logarithmic Returns (S&P 500, Nikkei 225)

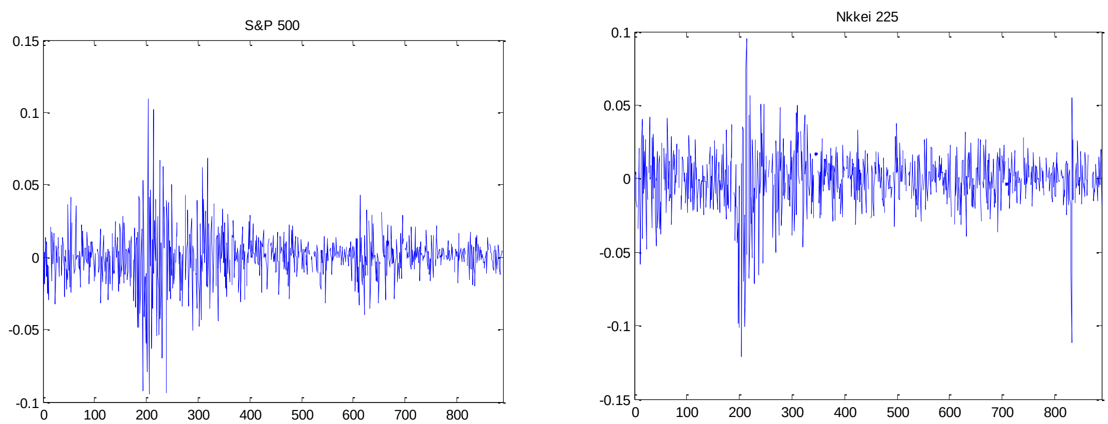

Figure 1 presents the plot of daily logarithmic returns for both series over the considered time period. We observe that volatility clustering is present in both cases as the two series show periods of low volatility which tend to be followed by periods of relatively low volatility and other periods of high volatility which likewise tend to be followed by high volatility. This aspect can be thought of as clustering of the variance of the error term over time, that is, the error term exhibits time varying heteroscedasticity.

4.3 Model Estimates

In this sub-section, first order EGARCH and GJR-GARCH models are fitted to the two empirical series and estimates obtained using the maximum likelihood estimation (MLE) and estimating function (EF) approaches. In parameter estimation under maximum likelihood method, we assume a standardized Gaussian or Student's – t distribution with $v = 10$ degrees of freedom for the innovations. Parameter estimates, corresponding standard errors (in parentheses), Akaike Information Criteria (AIC) and the log likelihood values are presented in *Tables 2 -6*.

Table 2: Parameter Estimates of EGARCH $(1,1)$ – S&P 500

Estimates Method	κ	β	γ_1	γ_2
MLE*	-0.108576 (0.051021)	0.986451 (0.008837)	0.220176 (0.092262)	-0.271144 (0.061067)
MLE**	-0.203312 (0.045728)	0.976439 (0.005239)	0.138657 (0.034506)	-0.166382 (0.028401)
EF	-0.204886 (0.043279)	0.976205 (0.004959)	0.138098 (0.032848)	-0.161534 (0.026946)

***Standardized Gaussian distribution ****Student's – t distribution ($v = 10$)

Table 3: Parameter estimates of GJR – GARCH $(1,1)$ - S&P500

Estimates Method	κ	α	β	γ
MLE*	1.96653E-06 (1.34835E-06)	0.00207731 (0.057016)	0.889019 (0.038039)	0.201962 (0.068537)
MLE**	2.36413E-06 (1.24408E-06)	0.00114628 (0.020089)	0.901695 (0.0205511)	0.173676 (0.0367019)
EF	2.37614E-06 (1.19813E-06)	0.00113883 (0.019185)	0.902273 (0.0196927)	0.16956 (0.0342864)

*****Standardized Gaussian distribution ******Student's – t distribution ($\nu = 10$)

Table 4: Parameter Estimates of EGARCH $(1,1)$ - Nikkei 225

Estimates Method	κ	β	γ_1	γ_2
MLE*	-0.232894 (0.100455)	0.98245 (0.010966)	0.201285 0.0918981	-0.199053 (0.048771)
MLE**	-0.280486 (0.072460)	0.966935 (0.008745)	0.140868 (0.039929)	-0.147923 (0.025432)
EF	-0.285148 (0.069517)	0.966433 (0.008371)	0.147311 (0.037503)	-0.147345 (0.023122)

*****Standardized Gaussian distribution ******Student's – t distribution ($\nu = 10$)

Table 5: Parameter estimates of GJR – GARCH $(1,1)$ – Nikkei 225

Estimates Method	κ	α	β	γ
MLE*	1.24593E-05 (4.05034E-06)	0.00125842 (0.061365)	0.860161 (0.030681)	0.21681 (0.072392)
MLE**	1.08575E-05 (3.23021E-06)	0.00379292 (0.025691)	0.859542 (0.027509)	0.188191 (0.041084)
EF	1.0501E-05 (2.96121E-06)	0.00618129 (0.023774)	0.857417 (0.026016)	0.185744 (0.036715)

*****Standardized Gaussian distribution ******Student's – t distribution ($\nu = 10$)

Table 6: AIC and Log – Likelihood Values

MODEL	**EGARCH $(1,1)$**		**GJR – GARCH $(1,1)$**	
SERIES	***S&P 500***	***Nikkei 225***	***S&P 500***	***Nikkei 225***
AIC	-6.0509	-6.3642	- 6.0463	-6.3637
Log - Likelihood	5538.71	6084.26	5532.15	6082.83

4.4 Discussion

From the results, it is seen that both models have almost similar AIC and Log – likelihood values for the two financial series. However EGARCH (1,1) has relatively higher Log – likelihood and lower AIC values than GJR – GARCH (1,1) indicating that it performs relatively better in explaining conditional volatility in both empirical series over the considered time period.

The coefficients γ_2 and γ for the first order EGARCH and GJR – GARCH models respectively reflects the leverage effects. The estimates indicate the magnitude and sign of the leverage effects. The EGARCH model shows a negative parameter of asymmetry in both financial series suggesting that past negative shocks (bad news) have a greater impact on subsequent volatility of returns than positive shocks (good news) do. The GJR – GARCH model records positive leverage effects, attesting that bad news in the market lead to a higher volatility of asset returns than good news.

From our parameter estimates it is clear that the EF approach is more efficient than the MLE method in parameter estimation of the first order EGARCH and GJR – GARCH models in finite samples.

The standard errors of the EF approach estimates are smaller than those of the maximum likelihood estimates assuming either a Gaussian or a Student's – t ($v = 10$) error distribution. The gain in efficiency follows from the fact that the EF approach does not rely on distributional specifications for optimality and that it accounts for higher order moments present in non – normal data such as most empirical financial time series. However, it is evident that the MLE method when assuming a Student's – t error distribution competes reasonably well with the EF approach and provides a better in-sample-fit than the MLE method when assuming a Gaussian error distribution across both data sets. This result is expected considering the Jarque –Bera normality test in *Table 1* which implies that the empirical distributions of the two return series exhibit heavier tails than the standard normal distribution. A Student's – t distribution exhibits excess kurtosis and fat tail behaviour.

5.0 Conclusion

In this paper our main goal was to derive optimal estimating functions for the Asymmetric GARCH modes in general and demonstrate the application of the EF approach as an alternative to the MLE approach in parameter estimation. We have shown that the EF approach competes reasonably well with the MLE method especially in cases where there are serious departures from normality in finite samples. This approach therefore provides a useful alternative method of estimation to the MLE method for the Asymmetric GARCH models especially in cases where the true distribution of the data is unknown as it does not rely on distributional assumptions for optimality. Extending the EF approach to the multivariate GARCH model is a subject for future research.

Acknowledgement

This research paper was prepared and made possible through the help and support of my academic supervisors; Prof. A.S Islam and Dr. L.A Orawo and The German Academic Exchange Service (*DAAD)* which sponsored my postgraduate studies.

Appendix (Proofs of Equations)

Proof of equation (14)

$$E\left(\lambda_{1t}\ \lambda_{2t}^{*}\right) \neq 0$$

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

Proof of equation (19)

$$E\left(\lambda_{1t}\ \lambda_{2t}\right) = E\left\{\left[y_t - x_t\omega\right]\left[\left(y_t - x_t\omega\right)^2 - h_t - \varepsilon_t\ \phi_{1t}\ h_t^{1/2}\right] / \psi_{t-1}\right\}$$

Since $E\left\{h_t\left(y_t - x_t\omega\right) / \psi_{t-1}\right\} = 0$,

$$\begin{aligned} E\left(\lambda_{1t}\ \lambda_{2t}\right) &= E\left\{\left[\left(y_t - x_t\omega\right)^3 - \left(y_t - x_t\omega\right)^2 \phi_{1t}\ h_t^{1/2}\right] / \psi_{t-1}\right\} \\ &= E\left\{\left(y_t - x_t\omega\right)^3 / \psi_{t-1} - \phi_{1t}\ h_t^{3/2}\right\} \\ &= \phi_{1t}\ h_t^{3/2} - \phi_{1t}\ h_t^{3/2} = 0 \end{aligned}$$

References

Fama, E. F. (1965). The Behaviour of Stock-Market Prices, Journal of Business. **38:** 34-105.
Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, Journal of Finance. **25:** 383-417.
Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica. **50:** 987-1007.
Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, Journal of Econometrics. **31:** 307-328.
Neelabh, R. (2009). Conditional Heteroscedastic Time Series Models with Asymmetry and Structural Breaks. Phd Thesis, University of Pune, India.
Nelson, D.B. (1991). Conditional heteroskedasticity in Asset returns: A new Approach, Econometrica. **59:** 347-370.
Glosten, L. R., Jagannathan, R. and Runkle, D. (1993). On the relation between the expected value and the volatility of nominal excess returns on stocks, Journal of Finance. **48:** 1779-1801.
Engle, R.F. and Ng V.K. (1993). Measuring and testing the impact of news on volatility, Journal of Finance. **48 (5):** 1749-1778.
Ding, Z., Granger, C. and Engle, R. (1993). A long memory property of stock returns and a new model, Journal of Empirical Finance. **1:** 83–106.
Zakoian, J.M. (1994). Threshold Heteroskedasticity Models, Journal of Economic Dynamics and Control. **19:** 931-944.
Sentana, E. (1995). Quadratic ARCH Models, Review of Economic Studies.**62:**639-661.
Godambe,V.P. (1960). An optimum property of regular maximum likelihood equations, Annals of Mathematical Statistics. **31:** 1208-1211.
Godambe, V.P. and Thompson, M.E. (1989). An extension of the Quasi-likelihood Estimation, Journal of Statistical Planning and Inference.**22:** 137-172.
Black, F. (1976). Studies of stock price volatility changes, Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, American Statistical Association, pp. 177-181.
Godambe, V.P. (1985). The foundations of finite sample estimation in Stochastic processes, Biometrika. **72:** 319 - 328.
Hyde, C.C. (1997). Quasi-Likelihood and Its applications, New York: Springer- Verlag
Zivot. E. (2008). Practical Issues in the Analysis of Univariate GARCH Models, Handbook of Financial Time Series, Springer, New York.