Research Journal of Applied Sciences 5 (2): 108-111, 2010
ISSN: 1815-932X

Aperformance Ratings of an Autocovariance Base Estimator (ABE) in the Estimation of GARCH Model Parameters When the Normality Assumption is Invalid

Daniel Eni
Department of Mathematics and Computer Science,
Federal University of Petroleum, Effurun, Delta State-Nigeria

Abstract: In this study, the performance of an Autocovariance Base Estimator (ABE) for GARCH models was studied, against that of the Maximum Likelihood Estimator (MLE) if the distribution assumption is wrongly specified as normal. We do this by first simulating time series data that fits GARCH model using the Log normal and t-distribution with degrees of freedom of 5, 10 and 15 as the true probability distribution but assumed normality in the process of parameter estimations. To keep track of consistency, we conduct and present the studies in sample sizes of 200, 500, 1000 and 1200. The two methods were then used to analyse the series under normality assumption. The result shows that the ABE method appears to be competitive in the situations considered.

Key words: Autocovariance functions, parameter estimation, normality, probabibilty, distribution GARCH, invalid

INTRODUCTION

The assumption of constant variance in the traditional time series models of ARMA is a major impediment to their applications in financial time series data where heteroscedasticity is obvious and cannot be neglected. To solve the stated problem, Engle (1982) proposed Autoregressive Conditional Heteroscedascity (ARCH) model. However, Engle in his first application of ARCH noted that a high order of ARCH is needed to satisfactorily model time varying variances. It is noted that many parameters in ARCH will create convergence problems for maximization routines. To avoid these problems, Bollerslev (1986) extended Engle's model to Generalized Autoregressive Conditional. Heteroscedasticity models (GARCH).

This models time-varying variances as a linear function of past square residuals and of its past value. It has proved useful in interpreting volatility clustering effects and has wide acceptance in measuring the volatility of financial markets. The ARCH and GARCH models are known as symmetric models. Other extensions based on observed characteristic of financial time series data are:

- The asymmetric models of which the exponential GARCH (EGARCH) model of Nelson (1991), the model of Gosten *et al.* (1993) (GJR-GARCH) of as well as the threshold model (T GARCH) of Zakoian (1994) are representatives models. These modes and interpret leverage effect where volatility is negatively correlated with returns
- The Fractionally Integrated GARCH model (FIGARCH) of Baillie *et al.* (1996) introduced to model long memory via the fractional operator $(1-L)^d$
- The GARCH in mean models that allows the mean to influence the variance

These models are popularly estimated by the Quasi-maximum Likelihood Method (QMLE) under the assumption that the distribution of one observation conditionally to the past is normal. The asymptotic properties of the estimator are well established. Weiss (1986) showed that the QMLE estimates are consistent and asymptotically normal under the fourth moment conditions.

These were again proved by Ling and Mcaleer (2003), under only the second moment conditions. If the assumption of normality is satisfied by the data then the method will produce efficient estimates otherwise inefficient estimates will be produced. Engle and Gonzalez-Revera (1991) studied the loss of estimation efficiency inherent in QMLE and concluded it may be severe if the distribution density is heavy tailed.

The QMLE estimator requires the use of numerical optimization procedure which depends on different optimization techniques for implementation. This potentially leads to different estimates. This is confirmed by recent studies by Brooks *et al.* (2001) and McCullough and Renfro (1999). Both reported different QMLE estimates across various packages using different optimization routines. These techniques estimates time varying variances in different ways and may result to different interpretations and predictions with varying implications to the economy. It is therefore important to undertake studies that would develop appropriate techniques for estimating parameters of processes used in modeling time series data.

To solve the stated problems, Eni and Etuk (2006) developed an Autocovariance Base Estimator (ABE) for estimating the parameters of GARCH models through an ARMA transformation of the GARCH model equation.The thrust of this study is to rate the performance of the Autocovariance Base Estimator when the normality assumption is violated.

The Autocovariance Base Estimator (ABE): Consider the GARCH (p, q) equation:

$$h_t = w_0 + \sum_{i=1}^{p} \alpha_i e_{t-1}^2 + \sum_{j=1}^{q} B_j h_{t-1} \quad (1)$$

Or its ARMA (Max (p, q),q) transform:

$$\varepsilon_t^2 = w_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + B_i)\varepsilon_{t-1}^2 - \sum_{j=1}^{q} B_j a_{t-1} + B_0 a_t \quad (2)$$

$$\varepsilon \sim N(0, \sigma^2)$$

To obtain the autoregressive parameters, we take advantage of the fact that the variance, var $(\varepsilon_t^2\, \varepsilon_{t-1}^2)$ for i>q in Eq. 2 will contain no moving average parameter B_1. Hence we set i = q + 1 … q + p to get the estimator:

$$\begin{bmatrix} V_{q+1} \\ V_{q+2} \\ \vdots \\ V_{q+\max(p,q)} \end{bmatrix} = \begin{bmatrix} Vq & V_{q-1} & \cdots & V_{q-(p-1)} \\ V_{q-1} & V_q & \cdots & V_{q-(p-2)} \\ \vdots & \vdots & \vdots & \\ V_{q-(p-1)} & V_{q-(p-2)} & \cdots & Vq \end{bmatrix} \begin{bmatrix} (\alpha_1 + B_1) \\ (\alpha_2 + B_2) \\ \vdots \\ (\alpha_p + Bq) \end{bmatrix} \quad (3)$$

Where V_i is the set of variances associated with Eq. 2. We can easily obtain the autoregressive parameters $\alpha_i + B_i$ by solving Eq. 3. Eni and Etuk (2006) have shown that the moving average parameters can be obtained from:

$$\sum_{i=0}^{p} f(\Phi_i) \begin{vmatrix} V_i & V_{i-1} & \cdots & V_{I-p} \\ V_{i+1} & V_i & \cdots & V_{i-(p-1)} \\ \vdots & \vdots & \vdots & \vdots \\ V_{i+q} & V_{i-q-1} & \cdots & V_{i+q-p} \end{vmatrix} \begin{bmatrix} \Phi_0 \\ -\Phi_1 \\ \vdots \\ -\Phi_p \end{bmatrix} = \sigma_a^2 \begin{bmatrix} B_0 & -B_1 & \cdots & -B_{q-1} & -B_q \\ -B_1 & -B_2 & \cdots & -B_q & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -B_{q-1} & -B_q & \cdots & 0 & 0 \\ -B_q & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} B_0 \\ -B_1 \\ \vdots \\ -B_{q-1} \\ -B_q \end{bmatrix} \quad (4)$$

or

$$\sum_{i=0}^{p} f(\Phi_i) V\Phi = \sigma_a^2 B\, b \quad (5)$$

where, f (Φ_i = -Φ_i, f (Φ_0) = Φ_0, Φ_i = (a_i + B_i)

Note that the quantity $\sum_{i=0}^{p} f(\Phi_i) V\Phi$ is already known. The variance V having been calculated from the data and the autoregressive parameters having been calculated from Eq. 3.

We find the moving average parameters Bi by solving the system:

$$F(B) = \sum_{i=0}^{p} f(\Phi_i) V\Phi - \sigma_r^2 B\, b = 0 \quad (6)$$

Eqution 5 is nonlinear and the solution can be found only through an iterative method. A ready procedure to consider is the one which depends on the Newton-Raphson algorithm. In this case, the B_{r+1} solution is obtained from the rth approximation according to:

$$B_{r+1} = B_r - \{f'(B_r)\}^{-1} f(B_r) \quad (7)$$

where, f (B_r) and f' (B_r) represent the vector Eq. 5 and its derivative evaluated at B = B_r. We note that:

$$f'(B) = \sigma_a^2 \begin{bmatrix} B_0 & B_1 & \cdots & B_{q-1} & B_q \\ -B_1 & B_2 & \cdots & B_q & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ -B_{q-1} & B_q & \cdots & 0 & 0 \\ -B_q & 0 & \cdots & 0 & 0 \end{bmatrix} + \sigma_a^2 \begin{bmatrix} B_0 & B_1 & \cdots & B_{q-1} & B_q \\ 0 & -B_0 & \cdots & B_{q-z} & B_{q-1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & -B_0 & B_1 \\ 0 & 0 & \cdots & 0 & -B_0 \end{bmatrix}$$

$$= \boldsymbol{\sigma}_r^2 (D_1 + D_2)$$

So that (4.12) becomes:

$$B_{r+1} = B_r + \{\sigma_a^2 (D_1 + D_2)\}^{-1} \left[\sum_{i=0}^{p} f(\Phi_1) V\,\Phi - \sigma_a^2 B\, b \right] \quad (8)$$

The starting point for the iteration Eq. 8 is $\boldsymbol{\sigma}_r^2 = 1$, $B_0 = V_0$, $B_i = 0$, i = 1 … q

Having computed the Autoregressive parameters Φ_i = (α_i + B_i) and the Moving average parameter B_i, it is easy to obtain the GARCH (p, q) parameters α_i and the constant parameter w_0 which is estimated using:

$$W = E(\varepsilon_t^2)(1 - \sum_{i=1}^{p} \alpha_i - \sum_{i=1}^{B} B_i) \quad (9)$$

MATERIALS AND METHODS

In this study, the Data Generating Process (DGP) involves the simulation of 1,500 data points with 10 replications using the random number generator in MATLAB 5. The random number generator in MATLAB 5 can generate all the floating point numbers in the interval [2^{-53}, 1-2^{-53}]. Hence it can generate 2^{1492} values before repeating itself. We note that the data points of

1,500 are equivalent to $2^{10.55}$ and with 10 replications; we will have a mere $2^{13.87267}$ data points. Hence the 1,500 data points with 10 replications were obtained without repetitions. Also, we used a program implementation for ARMA due to McLeod and Sales (1983) to find the QMLE. Although, we would assume Normality, we actually simulated the data points using Log normal distribution and the t-distribution with degree of freedom of 5,10 and 15. Of the 1,500 data points generated for each of the process, the first 200 observations were discarded to avoid initialization effects, yielding a sample size of 12000 observations. The results are reported in sample sizes of 200, 500, 1000 and 1,200.

These sample presentations are to enable us keep track of consistency and efficiency of the estimators. The relative efficiency of the Autocovariances Based Estimator (ABE) and the Quasi-maximum Likelihood (QML) estimators were studied under this misspecification of distribution function. The selection criteria used is the Aikake Information Criteria (AIC). For simulating the data points, the conditional variance equation for low persistence due to Engle and Ng (1993) is adopted:

$$h_t = 0.2 + 0.05\,\varepsilon_{t-1}^2 + 0.75h_{t-1}$$
$$\varepsilon_{t-1}^2 = h_t Z_t^2$$

and Z^2_t is any of

$$Z \sim t^5 \text{ or } Z \sim t_{10} \text{ or } Z \sim LN\,(0, 1) \text{ or } Z \sim t_{15}$$

where, N = normality, t_V = t-distribution with V degree of freedom, LN = Log normal.

RESULTS AND DISCUSSION

Apart from the parameter setting in the DGP, selected studies of the paramete settings (W, α, B) = (0.1, 0.15, 0.85) and (W, α, B) = (0.1, 0.25, 0.65) due to Lumsdain (1995) as well (W, α, B) = (1, 0.3, 0.6) and (W, α, B) = (1, 0.05, 0.9) due to Yi-Ting (2002) where also studied and the results obtained are in agreement with the result obtained from detail studies of the DGP. The results obtained from the DGP are shown in Table 1.

Table 1 shows the result under a sample size of 200 data points. The Table 1 reveals that the estimates are poor for QMLE and ABE. However, on the bases of the Aikate Information Criteria (AIC), the QMLE performed better than the ABE except under log normal distribution where ABE performed better than the QMLE.

A study of Table 2 showed that the estimates are better although still poor. The performance bridge between QMLE and ABE is closing. This can be seen from the AIC of QMLE and ABE under the different probability distribution functions except in the case of the log normality. Here and surprisingly too the QMLE

Table 1: Performance rating of QMLE and ABE for sample size of 200

	Method of estimation							
	QMLE				ABE			
Estimates	W	α	B	AIC	W	α	B	AIC
T (5)	0.16	0.010	0.77	-70.90	1.140	0.016	0.74	-65.312
T (10)	0.14	0.014	0.76	-140.36	1.138	0.012	0.75	-124.31
T (15)	0.15	0.170	0.76	-169.40	1.420	0.016	0.76	-157.21
LN (0,1)	9.30	-0.200	0.86	129.17	6.200	0.200	0.81	108.23

Table 2: Performance rating of QMLE and ABE Estimates under a sample size of 500

	Method of estimation							
	QMLE				ABE			
Estimates	W	α	B	AIC	W	2	α	AIC
T (5)	0.115	0.020	0.739	-132.341	0.15	0.025	0.73	-151.24
T (10)	0.018	0.034	0.742	-1021.220	0.21	0.029	0.74	-956.31
T (15)	0.170	0.036	0.750	-1973.420	0.20	0.030	0.75	-1472.40
LN (0,1)	6.790	-0.150	0.880	289.390	3.94	0.080	0.80	108.21

Table 3: Performance rating of QMLE and ABE estimates under sample size of 1000

	Method of estimation							
	QMLE				ABE			
Estimates	W	α	B	A/C	W	α	B	A/C
T (5)	0.180	0.020	0.74	-137.12	0.19	0.03	0.73	-140.12
T (10)	0.193	0.029	0.75	-1141.62	0.22	0.03	0.74	-1094.72
T (15)	0.195	0.034	0.75	-1984.71	0.21	0.04	0.75	-1976.22
LN (0,1)	5.240	-0.400	0.86	119.72	3.50	0.07	0.79	101.13

Table 4: Performance rating of QMLE and ABE estimates under sample size of 1,200

	Method of estimation							
	QMLE				ABE			
Estimates	W	α	B	A/C	W	α	B	A/C
T (5)	0.150	0.030	0.760	-162.11	0.18	0.039	0.730	-173.70
T (10)	0.018	0.039	0.743	-1391.30	0.19	0.041	0.740	-1350.11
T (15)	0.190	0.043	0.746	-2441.30	0.19	0.044	0.742	-2430.39
LN (0,1)	4.300	0.080	0.810	168.59	3.48	0.060	0.780	256.23

method failed to show consistency. We also note that the performance of the two method are enhanced under the t-distribution as the degree of freedom increases. An examination of Table 3 shows that both estimation models that is the QMLE and the ABE have equal performance ratings. However, the ABE has an edge in its performance under $t_{(5)}$ and LN (0,1) while QMLE has an edge under $t_{(10)}$ and $t_{(15)}$. The estimates under $t_{(15)}$ and $t_{(10)}$ are close to their true values for both estimation methods. Generally, the two methods gave consistent estimates.

The result in Table 3 is further confirmed by examining Table 4 where the two methods have nearly equal rating judging from the values of their AIC. This is inspite of its very poor performance at the sample size of 200.

CONCLUSION

The study in this section shows that the ABE method is adequate in estimating GARCH model parameters and can perform as well as the maximum likelihood estimate for reasonable large data point when the distribution assumption is miss-specified.

REFERENCES

Baillie, R.T., T. Bollerslev and H.O. Mikkelson, 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. J. Econ., 74: 3-30.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroscedasticity. J. Econ., 31: 307-327.

Brooks, I., S. Burke and G. Perssand, 2001. Benchmarks and accuracy of the GARCH model estimation. Int. J. Forecasting, 17: 45-56.

Engle, F. and G. Gonzalez-Revera, 1991. Semi parametric ARCH models. J. Bus. Econ. Stat., 19: 3-29.

Engle, F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. Econometrica, 50: 987-1008.

Engle, R.F. and V.K. Ng, 1993. Measurement and testing the impact of news on volatility. J. Fin., 48: 1749-1778.

Eni, D. and E.H. Etuk, 2006. Parameter estimation of GARCH models: An autocovariance approach. Proceedings of International Conference on New Trends in the Mathematical and Computer Sciences with Applications to Real World Problems, (ICMCS'06), Covenant University Ota, Nigeria, pp: 357-368.

Gosten, L., R. Jaganatan and D. Runkle, 1993. On the relationship between the expected value and the volatility of the nominal excess return on stock. J. Fin., 48: 1779-1801.

Ling, S. and M. Mcaleer, 2003. Asymptotic theory for a vector GARCH (1.1) quasi-maximum likelihood estimator. Econ. Theory, 19: 280-310.

Lumsdain, R.L., 1995. Consistency and asymptotic normality of the quasi-maximum likelihood estimator. Econometrica, 64: 575-596.

McCullough, B.D. and C.G. Renfro, 1999. Bench marks and software standard: A case study of GARCH procedure. J. Econ. Social Measurement, 25: 59-71.

McLeod, A. and P. Sales, 1983. An algorithm for approximate likelihood calculation of ARMA and seasonal ARMA models. J. R. Soc. C: Applied Stat., 32: 211-223.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica, 59: 347-370.

Weiss, A., 1986. Asymptotic theory of GARCH (1.1) model. Econ. Theory, 2: 107-131.

Yi-Ting, C., 2002. On the robustness of ljung-box and mcleed-li Q test: A simulation study. Econ. Bull., 3: 1-10.

Zakoian, J.M., 1994. Threshold heteroskedastic models. J. Econ. Dyn. Control, 18: 931-955.