The Philippine Statistician, 2002
Vol. 51, Nos. 1-2, pp. 31-41

Estimating the I(d)-GARCH(p,q) in Mean Model

Alex C. Gonzaga*

received August, 2002; revised October, 2002

ABSTRACT

In this paper, we propose the I(d)-GARCH(p,q) in mean model. It incorporates the volatility model GARCH(p,q) into the fractionally differenced process I(d). This allows the conditional variance to affect the mean. Some properties of this process are derived and an approximate maximum likelihood estimation procedure is proposed.

KEYWORDS AND PHRASES: long-memory, fractional differencing, GARCH(p,q)

1. INTRODUCTION

Modeling volatility of economic time series is an important aspect in financial risk management. GARCH (Generalized autoregressive conditional heterokedasticity) models are the most popular in modeling volatility. They are able to effectively remove excess kurtosis in return series. Moreover, Gokcan (2000) demonstrated that for emerging stock markets, GARCH model performs better than the nonlinear EGARCH model even if the stock market return series display skewed distributions. Baillie, Bollerslev and Mikkelsen (1996), and Chung (2002) modified the GARCH model to incorporate the idea of long-memory fractional differencing into the volatility model. The resulting volatility model is called the fractionally integrated GARCH or the FIGARCH model.

In asset pricing, expected returns have been shown to be related to volatility. Risk-averse economic agents require time-varying premiums as reward for bearing financial risks. Hence, time series of asset prices must not only depend on their movement overtime but also include volatility, a measure of risk, as a determinant of price. This inadequacy of the traditional expectation to explain the observed data has been supported by a series of papers (Karanosos(2001)).

Engle, Lilien and Robins (1987) introduced the ARCH-M mean model, which allows the conditional variance to affect the mean. It is also called WN-ARCH(p) in mean model, which is a white-noise with ARCH-in-mean effects. It incorporates the volatility model ARCH(p) into a white-noise mean model . Karanasos (2001) extended this model to ARMA with GARCH-in-mean effects. It incorporates the volatility model GARCH(p,q) into the general ARMA(r,s) process. Optimal predictors and the covariance structure of the model were presented. These results, however, were of theoretical purity and not intended for practical purposes.

In this paper, we are concerned with the analysis of a long-memory process allowing volatility to affect the mean. We incorporate GARCH(p,q) volatility into the intermediate-memory or long-memory process I(d) and we call the model an I(d)-GARCH(p,q) in mean model. Changing the conditional variance directly affects the expected return. Unlike the FIGARCH model, I(d)-GARCH(p,q) need not exhibit long-memory in the variance. This

* Associate Professor, College of Arts and Sciences, University of the Philippines Manila,
Email Address: alexg@wanted.ph

model, for instance, implies that the determinant of asset price is not only the time series movement over time, but also its volatility or the associated financial risk. We present the moving average representation, the covariance structure and an estimation procedure, which may be applied for practical purposes.

This paper is organized as follows. In Sections 1 and 2, we give an introduction and we discuss preliminary concepts. In Sections 3, we introduce the model and we present the main results. We then give some concluding remarks in Section 4.

2. PRELIMINARY CONCEPTS

In this section, we present some concepts and standard results that will be used in the discussion of the main results.

An ARMA process $\{X_t\}$ is usually referred to as a *short-memory process* since the autocorrelation between X_t and X_{t+k} decreases rapidly at an exponential rate to zero as $k \to \infty$, that is, $\rho(k) \sim Cr^{-k}$, k = 1,2,.., where C>0 and 0<r<1. Brockwell (1987) defines a *long-memory process* as a stationary process for which $\rho(k) \sim Ck^{2d-1}$ as $k \to \infty$, where C>0 and d < 0.5. In this case, the autocorrelations decay to zero slowly at a hyperbolic rate. For our purpose, if d < 0 and $\sum_{k=-\infty}^{\infty} |\rho(k)| < \infty$, we call $\{X_t\}$ an *intermediate-memory process*. It is a *long memory* process when 0 < d < 0.5 and $\sum_{k=-\infty}^{\infty} |\rho(k)| = \infty$.

Long-memory processes are often modeled by means of the *fractionally integrated* I(d) process. (For our purpose, we say that a stochastic process is *stationary* if it is covariance stationary.) A *fractionally integrated* I(d) process $\{X_t\}$ is a stationary process such that

$$(1-L)^d X_t = \varepsilon_t$$

where ε_t is white noise , L is the backshift operator such that $LX_t = X_{t-1}$, $(1-L)^d$ is the fractional difference operator. If $d \in (0,0.5)$, $\{X_t\}$ is long-memory process (nonsummable autocorrelations). If $d \in (-0.5,0)$, $\{X_t\}$ is an intermediate-memory process (summable autocorrelations). The upper bound d<0.5 is needed, because for $d \geq 0.5$, the process is not stationary. However, the case d>0.5 can be reduced to the case -0.5<d<0.5 by taking appropriate integer differencing. For instance, if d = 1.4, then the differenced process $(1-L)^d W_t$ is the stationary solution of $(1-L)^d X_t = \varepsilon_t$ with d = 0.4 and $W_t = (1-L)X_t$.

Engle, et al (1982) proposed the ARCH (autoregressive conditional heteroskedasticity) model. It has been useful in explaining and forecasting volatility. The ARCH(q) model characterizes the distribution of the stochastic error ε_t conditional on the realized values of the set $I_{t-1} = \{\varepsilon_{t-1}, \ldots, \varepsilon_{t-q}\}$:

$$\varepsilon_t | I_{t-1} \sim N(0, h_t).$$

The ARCH(p) model can be formulated as

$$h_t = a_0 + a_1\varepsilon_{t-1}^2 + a_2\,\varepsilon_{t-2}^2 + \ldots + a_q\varepsilon_{t-q}^2,$$

where $a_0 > 0$ and $a_i \geq 0$, i=1,...,q, to ensure that the conditional variance is positive. An explicit generating equation for an ARCH process is

$$\varepsilon_t = \eta_t \sqrt{h_t}$$

where $\eta_t \sim NID(0,1)$. Clearly, ε_t is conditionally normal with mean zero and variance h_t.

Bollerslev (1986) proposed a generalization of the ARCH model, which he termed Generalized ARCH or GARCH. He suggested that the conditional variance be specified as

$$h_t = a_0 + a_1\varepsilon_{t-1}^2 + \ldots + a_q\varepsilon_{t-q}^2 + b_1h_{t-1} + \ldots + b_ph_{t-p}, \qquad (1)$$

where $a_0 > 0$, $a_i \geq 0$ for $i=1,\ldots,q$, $b_i \geq 0$ for $i=1,\ldots,p$ are imposed so that the conditional variance is strictly positive. Hence, the GARCH(p,q) specifies the conditional variance to be a linear combination of p lags of the squared residuals and q lags of the conditional variance h_t. In Equation (1), if we add and subtract a_ih_i, $i=1,2,\ldots,q$, we get

$$h_t = a_0 + a_1(\varepsilon_{t-1}^2 - h_{t-1}) + \ldots + a_q(\varepsilon_{t-q}^2 - h_{t-q}) + b_1^* h_{t-1} + \ldots + b_m^* h_{t-m},$$

where $m = \max(p,q)$, $a_i = 0$ for $i > q$, $b_i = 0$ for $i > p$, $b_i^* = b_i + a_i$, $i=1,2,\ldots,m$ and $v_t = \varepsilon_t^2 - h_t$. Clearly, v_t has mean zero and serially uncorrelated. Hence, it can be treated as an innovation. The ARMA representation of h_t is given by

$$B^*(L)\, h_t = A(L)v_t \qquad (2)$$

where $B^*(x) = 1 - \sum_{j=1}^{m}(a_j + b_j)x^j$ and $A(x) = 1 + \sum_{j=1}^{q} a_j x^j$.

The unconditional mean of the GARCH(p,q) model is given by

$$E(h_t) = a_0 + \sum_{i=1}^{q} a_i E(\varepsilon_{t-i}^2) + \sum_{i=1}^{p} b_i E(h_{t-i}).$$

Since $E(\varepsilon_t^2) = E(E(\varepsilon_t^2 \mid I_{t-1})) = E(h_t)$,

$$E(\varepsilon_t^2) = E(h_t) = \frac{a_0}{1 - \sum_{i=1}^{p} b_i - \sum_{i=1}^{q} a_i}. \qquad (3)$$

Hence, we impose the condition $\sum_{i=1}^{p} a_i + \sum_{i=1}^{q} b_i < 1$ for the existence of a finite variance of the innovation process $\{\varepsilon_t\}$. The unconditional variance of ε_t is constant, although the conditional variance, $E(\varepsilon_t^2 \mid I_{t-1}) = h_t$, changes with time.

3. MAIN RESULTS

In this section, we present some results that may be used in the analysis of the I(d)-GARCH(p,q) in mean model. We derive the moving average representation of y_t in terms of the innovations ε_t and v_t, which are shown to be uncorrelated. We also obtain the covariance of y_t and we show that y_t is an intermediate-memory or a long-memory process. Finally, we propose an approximate maximum likelihood estimation procedure of the I(d)-GARCH(p,q) parameters.

3.1 The Model

We define the I(d)-GARCH(p,q) in mean model by

$$(1-L)^d y_t = \delta h_t + \varepsilon_t \qquad (4)$$

where

$$h_t = a_0 + a_1\varepsilon_{t-1}^2 + a_2\,\varepsilon_{t-2}^2 + \ldots + a_q\varepsilon_{t-q}^2 + b_1 h_{t-1} + \ldots + b_p h_{t-p},$$
$$\varepsilon_t | I_{t-1} \sim N(0, h_t),$$

and $a_0 > 0$, $a_i \geq 0$ for i=1,...,q, $b_i \geq 0$ for i=1,...,p, $-0.5 < d < 0.5$. In the I(d)-GARCH(p.q) model, we allow the conditional variance to affect the mean. Hence, changing the conditional variance directly affects the expected return.

3.2 Moving Average Representation

In this section, we give the moving average representation of the I(d)-GARCH(p,q) process y_t. We express y_t as an infinite linear combination of innovations ε_t and v_t, which are shown to be uncorrelated. We assume that the stationarity and invertibility conditions are satisfied, that is, the roots of A(x) and B*(x) lie outside the unit circle. Moreover, we assume that A(x) and B*(x) have no common zeros.

From Equations (2) and (4), we have

$$y_t = (1-L)^{-d}\,\varepsilon_t + \delta\,\frac{A(L)(1-L)^{-d}}{B^*(L)}\,v_t$$
$$y_t = (1-L)^{-d}\,\varepsilon_t + \delta G(L)\,v_t \qquad (5)$$

where

$$G(L) = \frac{A(L)(1-L)^{-d}}{B^*(L)} = \sum_{j=0}^{\infty} g_j L^j, \qquad C(L) = \frac{A(L)}{B^*(L)} = \sum_{j=0}^{\infty} c_j L^j,$$

$$W(L) = (1-L)^{-d} = \sum_{j=0}^{\infty} \omega_j L^j, \quad \omega_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}$$

and $\Gamma(\bullet)$ is the gamma function. Now,

$$G(L) = C(L)W(L) = \left(\sum_{i=0}^{\infty} c_i L^i\right)\left(\sum_{j=0}^{\infty} \omega_j L^j\right) = \sum_{j=0}^{\infty}\left(\sum_{j=0}^{i} c_j \omega_{i-j}\right) L^j .$$

Hence, $g_i = \sum_{j=0}^{i} c_j \omega_{i-j}$, where by a similar argument to the derivation of the moving average coefficients of ARMA model (Brockwell and Davis(1987)), $c_j = a_j - \sum_{i=1}^{p} b_i^* c_{j-i}$ with $c_0 = 1$, $a_j = 0$ for j>m. Thus, the moving average representation of y_t is given by

$$y_t = \sum_{j=0}^{\infty}\left(\delta g_j v_{t-j} + \omega_j \varepsilon_{t-j}\right).$$

Now, $Cov(\varepsilon_s, v_t) = E(\varepsilon_s v_t) - E(\varepsilon_s)E(v_t)$, for $t,s \in Z^+$. If s=t, then

$$Cov(\varepsilon_s, v_t) = E(\varepsilon_s v_t) - E(\varepsilon_s)E(v_t) = E(\varepsilon_t(\varepsilon_t^2 - h_t)) = E(E(\varepsilon_t^3 | I_{t-1})) - E(h_t\, E(\varepsilon_t | I_{t-1})) = 0.$$

If s<t, then

$$Cov(\varepsilon_s, v_t) = E(\varepsilon_s v_t) - E(\varepsilon_s)E(v_t) = E(\varepsilon_s(\varepsilon_t^2 - h_t) = E(\,E(\varepsilon_s\varepsilon_t^2|I_{t-1}) - E(\varepsilon_s h_t)$$
$$= E(\varepsilon_s E(\varepsilon_t^2|I_{t-1})) - E(\varepsilon_s h_t) = E(\varepsilon_s h_t) - E(\varepsilon_s h_t) = 0.$$

If s>t, then

$$Cov(\varepsilon_s, v_t) = E(\varepsilon_s v_t) - E(\varepsilon_s)E(v_t) = E(\varepsilon_s(\varepsilon_t^2 - h_t) = E(\varepsilon_s\varepsilon_t^2) - E(\varepsilon_s h_t)$$
$$= E(E(\varepsilon_s\varepsilon_t^2|I_{s-1})) - E(E(\varepsilon_s h_t|I_{s-1})) = E(\varepsilon_t^2\, E(\varepsilon_s|I_{s-1})) - E(h_t E(\varepsilon_s|I_{s-1})) = 0.$$

Thus, in the moving average representation the innovations ε_t and v_s are uncorrelated for all $t,s \in Z^+$.

3.3 Covariance Structure

In this section, we derive the covariance of y_t . We show that it has hyperbolic decay; hence, an intermediate-memory or a long-memory process.

From Equation (5), we have

$$y_t = (1-L)^{-d}\varepsilon_t + \delta G(L)v_t$$

Hence,

$$Cov_k(y_t) = cov_k\,((1-L)^{-d}\varepsilon_t) + cov_k(\delta G(L)v_t) + cov_k([(1-L)^{-d}\varepsilon_t][\,\delta G(L)v_t]).$$

Since $E((1-L)^d\varepsilon_t) = E(E((1-L)^d\varepsilon_t)|I_{t-1})) = 0$,

$$Cov_k((1-L)^d\varepsilon_t) = E\,(Cov((1-L)^d\varepsilon_t|I_{t-1})).$$

From Beran (1994),

$$Cov_k\,((1-L)^{-d}\varepsilon_t) = E\left[\frac{\text{var}(\varepsilon_t\,|\,I_{t-1})(-1)^k\,\Gamma(1-2d)}{\Gamma(k-d+1)\Gamma(1-k-d)}\right]$$
$$= \frac{E(h_t)(-1)^k\,\Gamma(1-2d)}{\Gamma(k-d+1)\Gamma(1-k-d)} = E(h_t)\;\eta(k).$$

where $E(h_t)$ is defined by Equation (3). Similarly,

$$Cov_k((1-L)^{-d}v_t)) = var(v_t)\;\eta(k) = E(v_t^2)\;\eta(k) = 2E(h_t^2)\;\eta(k).$$

The second term is

$$Cov_k(\delta G(L)v_t) = Cov(\delta G(L)v_t) = \delta^2 Cov(C(L)(1-L)^{-d}v_t)$$
$$= 2E(h_t^2)\delta^2 \sum_{h=0}^{\infty}\sum_{j=0}^{\infty}\left(c_j c_{j+h}\eta(k-h)\right).$$

Since ε_t and v_s are uncorrelated for all $t,s \in Z^+$, then

$$Cov_k([(1-L)^{-d}\varepsilon_t][\,\delta G(L)v_t]) = 0.$$

Hence, the covariance of y_t is given by

$$Cov_k(y_t) = E(h_t)\;\eta(k) + 2E(h_t^2)\delta^2 \sum_{h=0}^{\infty}\sum_{j=0}^{\infty}\left(c_j c_{j+h}\eta(k-h)\right), \quad (6)$$

where $E(h_t)$ is defined by Equation (3). Taking k=0 in Equation (6), we have

$$Var(y_t) = \frac{E(h_t)\Gamma(1-2d)}{\{\Gamma(1-d)\}^2} + 2E(h_t^2)\delta^2 \sum_{h=0}^{\infty} \sum_{j=0}^{\infty} \left(c_j c_{j+h} \eta(-h)\right),$$

where c_j are the coefficients in the moving average representation of h_t. The general expression for $E(h_t^2)$ is quite complicated to be expressed in closed form. However, for GARCH(1,1), from Karanasos (2002) we have

$$E(h_t^2) = \frac{a_0^2(1+a_1+b_1)}{(1-a_1-b_1)(1-3a_1^2-b_1^2-2a_1b_1)} .$$

Note that $\eta^*(k) = \sum_{h=0}^{\infty} \sum_{j=0}^{\infty} \left(c_j c_{j+h} \eta(k-h)\right)$ may be considered as the autocovariance of an ARFIMA process with moving average coefficients c_j and error variance 1. From Beran (1994), as $k \to \infty$ there is a constant $C_1 > 0$ such that

$$\eta^*(k) \sim C_1 k^{2d-1}.$$

Moreover, Stirling's formula gives us

$$\eta(k) \sim C_2 k^{2d-1}$$

for some constant C_2. Since $E(h_t)$ and $E(h_t^2)$ are independent of k, as $k \to \infty$

$$Cov_k(y_t) \sim C_3 k^{2d-1},$$

for some constant C_3. Hence, I(d)-GARCH(p,q) is stationary and has autocovariance with hyperbolic decay. Thus, it is also an intermediate-memory or a long-memory process.

From the moving average representation of y_t, the autocovariance function can also be written as

$$y_t = \delta^2\sigma_\varepsilon^2 \sum_{j=0}^{\infty} g_j g_{j+k} + \sigma_v^2 \sum_{j=0}^{\infty} \omega_j \omega_{j+k} ,$$

where σ_ε and σ_v are the unconditional variances of the innovations ε_t and v_t. Since ε_t and v_t are uncorrelated, from Brockwell and Davis (1987), the autocovariance generating function is

$$A^*(z) = \delta^2\sigma_\varepsilon^2 G(z)G(z^{-1}) + \sigma_v^2 W(z) W(z^{-1}), \qquad r^{-1} < |z| < r$$

for some $r > 1$, where

$$G(L) = \frac{A(L)(1-L)^{-d}}{B^*(L)} = \sum_{j=0}^{\infty} g_j L^j , \quad W(L) = (1-L)^{-d} = \sum_{j=0}^{\infty} \omega_j L^j .$$

3.4 Estimation of Parameters

In this section, we present an approximate maximum likelihood estimation procedure for the I(d)-GARCH(p,q) in mean model. This method may be used also to estimate the long-memory parameter d of a pure I(d) model by setting all other parameters zero.

From Equation (4), we have

$$(1-L)^d y_t = \delta h_t + \varepsilon_t$$

where $\varepsilon_t|I_{t-1} \sim N(0,h_t)$, $(1-L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} L^k$, and $\Gamma(\bullet)$ is the gamma function. Let $\Theta = (d, \delta, a_0, a_1, \ldots, a_q, b_1, b_2, \ldots, b_p)$ be the vector of all the parameters to be estimated. Hence, for a time series with realization $\{y_1, y_2, \ldots, y_T\}$, the approximate loglikelihood function is given by

$$\ln L(\Theta) = \frac{-T}{2} \ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T} \ln(h_t) - \frac{1}{2}\sum_{t=1}^{T}\left(\frac{\varepsilon_t^2}{h_t}\right).$$

The computation of the approximate maximum likelihood estimates may be carried out as follows. With the given values of the parameter vector $\Theta = (d, \delta, a_0, a_1, \ldots, a_q, b_1, b_2, \ldots, b_p)$, we compute the following recursively:

STEP 1. $\xi_t = \sum_{j=0}^{t-1} \omega_j(d)\left(\frac{y_{t-j} - \bar{y}}{s}\right)$, where $\bar{y}$ and s are the sample mean and standard deviation, respectively, of y_t and

$$\omega_k(d) = \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} = \prod_{0<n\leq k} \frac{n-1-d}{n}, \qquad k = 0,1,2,3\ldots$$

STEP 2. Compute h_t and ε_t alternately, that is, compute h_1, then ε_1, then h_2, then ε_2, and so on, where

$$h_t = a_0 + \sum_{j=1}^{q} a_j(\varepsilon_{t-j})^2 + \sum_{j=1}^{p} b_j h_{t-j}$$

$$\varepsilon_t = \xi_t - \delta h_t$$

STEP 3. Compute $\sum_{t=1}^{T}\left(\ln(h_t) + \frac{\varepsilon_t^2}{h_t}\right) = R(\Theta)$.

STEP 4. Choose Θ that minimizes $R(\Theta)$.

The pre-sample values of ε_t for t = 0, -1,-2, ... are all equal to 0, while the values of h_t for t = 0, -1, -2,... are all equal to 1. These pre-sample values are natural consequence to the fact that ε_t is assumed to have mean zero and h_t is the conditional variance of the error term; hence, the sample variance is a natural estimator. The choice of the sample size T and pre-sample values may be the contentious issues in the analysis.

The approximate maximum likelihood estimates may also be computed from the system of nonlinear equations involving the partial derivatives with respect to the parameters of the log-likelihood function.

Let $\Theta = (d, \delta, a_0, a_1, \ldots, a_q, b_1, b_2, \ldots, b_p)$, $\eta = (a_0, a_1, \ldots, a_q, b_1, b_2, \ldots, b_p)$,

$$e_t = \sum_{k=0}^{t-1} \omega_k(d)(y_{t-k} - \delta h_{t-k}),$$

$$S(\Theta) = \sum_{t=1}^{T}\left[ln(h_t(\eta)) + \frac{e_t^2(\Theta)}{h_t(\eta)}\right],$$

$$s_t(\Theta) = ln(h_t(\eta)) + \frac{e_t^2(\Theta)}{h_t(\eta)}.$$

The approximate maximum likelihood estimate minimizes $S(\Theta)$ with respect to Θ. This amounts to solving the system of q+p+3 nonlinear equations

$$\sum_{t=1}^{T} \dot{s}_t(\Theta) = 0,$$

where

$$\dot{s}_t(\Theta) = \left(\frac{\partial s_t(\Theta)}{\partial d}, \frac{\partial s_t(\Theta)}{\partial \delta}, \ldots, \frac{\partial s_t(\Theta)}{\partial b_p}\right)^T.$$

Note that

$$\frac{\partial s_t(\Theta)}{\partial d} = \frac{2e_t(\Theta)}{h_t(\eta)}\frac{\partial}{\partial d}e_t(\Theta),$$

$$\frac{\partial s_t(\Theta)}{\partial \delta} = \frac{2e_t(\Theta)}{h_t(\eta)}\frac{\partial}{\partial \delta}e_t(\Theta),$$

$$\frac{\partial s_t(\Theta)}{\partial a_i} = \frac{\varepsilon_{t-i}^2}{h_t(\eta)} + \frac{2e_t(\Theta)}{h_t(\eta)}\frac{\partial}{\partial a_i}e_t(\Theta) - \frac{e_t^2(\Theta)\varepsilon_{t-i}^2}{h_t^2(\eta)},$$

$$\frac{\partial s_t(\Theta)}{\partial b_j} = \frac{h_{t-j}}{h_t(\eta)} + \frac{2e_t(\Theta)}{h_t(\eta)}\frac{\partial}{\partial b_j}e_t(\Theta) - \frac{e_t^2(\Theta)h_{t-j}(\eta)}{h_t^2(\eta)},$$

where

$$\frac{\partial e_t(\Theta)}{\partial d} = \sum_{k=0}^{t-1} \omega_k(d)[\psi(k-d) - \psi(-d)](y_{t-k} - \delta h_{t-k}(\eta)),$$

where $\psi(\bullet)$ is the digamma function,

$$\frac{\partial e_t(\Theta)}{\partial \delta} = -\sum_{k=0}^{t-1} \omega_k(d)h_{t-k}(\eta),$$

$$\frac{\partial e_t(\Theta)}{\partial a_i} = -\sum_{k=0}^{t-1} \delta\varepsilon_{t-k-i}, \quad 0 \le i \le q,$$

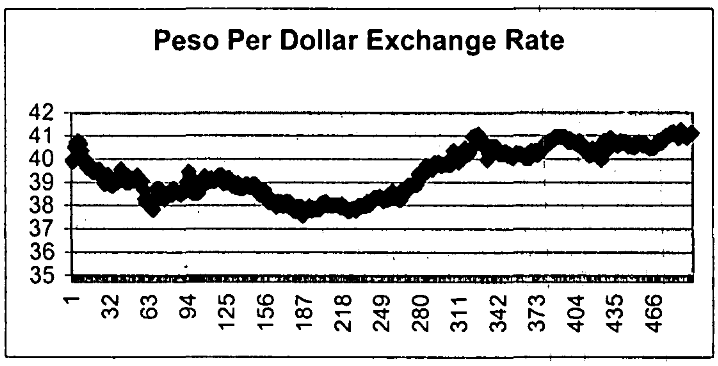
$$\frac{\partial e_t(\Theta)}{\partial b_j} = -\sum_{k=0}^{t-1} \delta h_{t-k-j}(\eta), \quad 1 \le j \le p.$$

These partial derivatives may also help significantly in the derivation of the asymptotic properties of the estimators.

3.5 Application

To illustrate the application of the estimation procedure described in the previous section, we analyze a data set consisting of daily observations of the exchange rate of peso per US dollar from September 23, 1999 to August 16, 2001.

Figure 1 Daily Peso Per Dollar Exchange Rate (X_t)



Preliminary analysis showed that the given series, x_t, does not represent a long-memory process. Aggregation of the data is implemented to artificially induce long memory. We consider the series

$$z_t = (x_{5t} + x_{5t+1} + x_{5t+2} + x_{5t+3} + x_{5t+4})/5.$$

The series z_t represents the weekly average peso per dollar value. Initial estimate of the long-memory parameter d gives a fractional value greater than 0.5. This indicates that the series is not stationary. Hence, the differenced series $y_t = z_t - z_{t-1}$ is obtained and used in the analysis. An I(d)-GARCH(1,1) is used to model the series y_t. Simultaneous estimation of the parameters d, a_0, a_1, b_1 and δ implementing Steps 1 to 4 is performed using S-PLUS 2000 Professional. Results showed the following estimates of the parameters:

$$d = 0.10,\ a_0 = 0.45,\ a_1 = 0.10,\ b_1 = 0.40,\ \delta = 0.10.$$

Hence, the I(d)-GARCH(1,1) model for y_t is given by

$$(1\text{-}L)^{0.10} y_t = 0.10\,(0.45 + 0.10\varepsilon_{t-1}^2 + 0.40h_{t-1}) + \varepsilon_t.$$

The corresponding volatility model is given by

$$h_t = 0.45 + 0.10\varepsilon_{t-1}^2 + 0.40h_{t-1}.$$

Similar procedure allows us fit a pure I(d) model by setting δ and the volatility parameters equal to zero. The I(d) model for y_t is given by

$$(1\text{-}L)^{0.0875} y_t = \varepsilon_t.$$

These estimates show that the value of the long-memory parameter is dependent on the values of δ and the volatility parameters. Clearly, the expected return of y_t of the I(d)-GARCH(1,1) model is nonzero; while that of the pure I(d) model is zero. Further, analysis comparing I(d)-GARCH(p,q) model and other long-memory models is currently in preparation.

4. CONCLUSIONS

In this paper, we presented some results useful in the analysis of the I(d)-GARCH(p,q) in mean model. Further research may be done on the asymptotic properties of the approximate likelihood estimators, the best linear predictor for forecasting of future values and model selection criterion such as AIC for the given model. Some contentious issues in computing the approximate maximum likelihood estimates, such as the choice of T and the values assumed in the pre-samples, may be resolved.

ACKNOWLEDGMENT

The author would like to thank the editors of The Philippine Statistician and the anonymous referee for their helpful and constructive comments and suggestions. Thanks also to Mr. Derick Erl Sumalapao for providing the data set and Mr. Richard Bryan Chua for sharing his programming expertise.

References

BAILLIE, R., BOLLERSLEV, T., AND MIKKELSEN, H.O. (1996) "Fractionally Integrated Generalized Autoregressive Conditional Heterokedasticity", *Journal of Econometrics* 74, pp. 3-30

BERAN, J. (1994). *Statistics for Long-memory Processes,* Chapman and Hall, New York.

BOLLERSLEV, T. (1986) "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics* 31, pp. 5-59.

BINSOL, A., BONZO, D. (1996). On the Application of ARFIMA Model to the Philippine Stock Exchange Oil Index, *The Philippine Statistician* Vol. 44-45, pp. 33-48.

BROCKWELL, P. AND DAVIS, R. (1987). *Time Series: Theory and Methods,* Springer-Verlag, New York.

CHUNG, C. (2002) "Estimating the Fractionally Integrated GARCH Model", Working Paper, *Academia Sinica.*

ENGLE, F.R. (1982) "Autoregressive Conditional Heteroskedasticity with Estimates of Variance of United Kingdom Inflation", *Econometrica* 50 (4), pp. 987-1008.

ENGLE, F.R., LILIEN, D.M. AND ROBINS, R.P. (1987) "Estimating Time Varying Premia in the Term Structure: the ARCH-M Model", *Econometrica* 55(2), pp. 391-407

GONZAGA, A. AND BONZO, D. (2000) "Wavet-based Exact Maximum Likelihood Estimation of ARFIMA(p,d,q) Parameters", *The Philippine Statistician*, 49(1-4), pp. 57-71

GONZAGA, A. AND BONZO, D. (2001a) "Wavelet-based Inference for Long-memory Process", *The Philippine Statistician*, 50(1-4), pp. 25-40

GONZAGA, A. AND BONZO, D. (2001b) "Inference for Long-memory Processes Using Local Lyapunov Exponents", *Manila Journal of Science* Vol. IV No. 2, pp. 5-15

GOKCAN, S. (2000) "Forecasting Volatility of Emerging Stock Markets: Linear versus Nonlinear GARCH Models", *Journal of Forecasting* 19, pp. 499-504

GRANGER, C.W.J.(1966). "The Typical Spectral Shape of an Economic Variable", *Econometrica* 34, pp. 150-161.

KARANASOS, M. (2001) "Prediction in ARMA Models with GARCH in Mean Effects", *Journal of Time Series Analysis* 22(5), pp. 555-576

KARANASOS, M. (2002) Lecture Notes in Macro and Financial Econometrics, Working Paper.

LeBARON, B. (1992) "Some Relations Between Volatility and Serial Correlations in Stock Returns", *Journal of Business* 65, pp. 199-219

PALMA, W. AND CHAN, A.N.H. (1997). "Estimation and Forecasting of Long-memory Processes with Missing Values", *Journal of Forecasting* Vol. 16, pp. 385-410.