A Range-Based GARCH Model for Forecasting Volatility

Mapa, Dennis S.

School of Statistics University of the Philippines Diliman, School of Economics, University of the Philippines Diliman

December 2003

A Range-Based GARCH Model for Forecasting Volatility

Dennis S. Mapa[1]

ABSTRACT

A new variant of the ARCH class of models for forecasting the conditional variance, to be called the Generalized AutoRegressive Conditional Heteroskedasticity Parkinson Range (GARCH-PARK-R) Model, is proposed. The GARCH-PARK-R model, utilizing the extreme values, is a good alternative to the Realized Volatility that requires a large amount of intra-daily data, which remain relatively costly and are not readily available. The estimates of the GARCH-PARK-R model are derived using the Quasi-Maximum Likelihood Estimation (QMLE). The results suggest that the GARCH-PARK-R model is a good middle ground between intra-daily models, such as the Realized Volatility and inter-daily models, such as the ARCH class. The forecasting performance of the models is evaluated using the daily Philippine Peso-U.S. Dollar exchange rate from January 1997 to December 2003.

Key Words: Volatility, Parkinson Range, GARCH-PARK-R, QMLE

I. INTRODUCTION

Since the introduction of the seminal paper on AutoRegressive Conditional Heteroskedasticity (ARCH) process of Robert Engle in 1982, researches on financial econometrics have been dominated by extensions of the ARCH process. One particular difficulty experienced in evaluating the various ARCH-type of models is the fact that volatility is not directly measurable – the conditional variance is unobservable. The absence of such a benchmark that we can use to compare forecasts of the various models makes it difficult to identify good models from bad ones.

Anderson and Bollerslev (1998) introduced the concept of "realized volatility" from which evaluation of ARCH volatility models are to be made. Realized

---

[1] The author is an Assistant Professor of Statistics at the School of Statistics, University of the Philippines, Diliman and a Ph.D. (Candidate) at the School of Economics, U.P.Diliman. The author wish to thank Professors Adolfo M. de Guzman, Lisa Grace S. Bersales and Joselito C. Magadia for their comments and suggestions. All remaining errors are the authors' responsibility.

volatility models are calculated from high-frequency intra-daily data, rather than inter-daily data. In their seminal paper, Anderson and Bollerslev collected information on the DM-Dollar and Yen-Dollar spot exchange rates for every five-minute interval, resulting in a total of 288 5-minute observations per day! The 288 observations were then used to compute for the variance of the exchange rate of a particular day. Although volatility is an instantaneous phenomenon, the concept of realized volatility is by far the closest we have to a "model-free" measure of volatility.

Obviously, there is a trade-off when one is interested in estimating the conditional variance using realized volatility. While it may provide a model-free estimate of the unknown conditional variance, the data requirement (getting observation every 5 minutes, for instance) is simply enormous. In the case of the Philippines, the Philippine Stock Exchange (PSE) starts trading at 9:30 a.m. up to 12:00 noon, for a total of 150 minutes of trading time or 30 5-minute observations. Given the low market activity, it is highly probable that the price of a particular stock will not move during that 5-minute period. The same problem might be encountered in the foreign exchange market in the Philippine Dealing System (PDS). Data problem may hinder the use of realized volatility for emerging markets such as the Philippines.

An alternative approach to model volatility using intra-daily data is through the use of the range, the difference between the highest and lowest values for the day. The range is the popular measure of volatility (the standard deviation) in the area of quality control. The range is convenient to use, especially for researchers who do not have access to information on the trading floors of various markets, since major newspapers normally report the highest and lowest values of assets (stock prices, currencies, interest rates, etc.), together with the opening and closing prices.

This paper proposes the use of the Range, specifically the Parkinson Range, in estimating the conditional variance. The model will be called the Generalized AutoRegressive Conditional Heteroskedasticity Parkinson Range (GARCH-PARK-R) model. This paper is organized as follows: section 1 serves as the introduction, section 2 discusses the ARCH process and its extensions. Section 3 introduces the concept of realized volatility. The GARCH-PARK-R model and the estimation procedure are discussed in section 4. Section 5 provides the empirical results and section 6 concludes.

II. The ARCH Process and its Extensions

In this section, the ARCH process will be defined and some of its important properties discussed. Hopefully, doing it at this early stage will serve two purposes. First is to acquaint the readers of the ARCH process for them to fully appreciate the survey of the literature discussed. Secondly, for them to have a better understanding of the important properties of the ARCH process that made it very attractive in modeling financial time series.

Let $\{u_t(\theta),\ t \in (\ldots,-1,0,1,\ldots)\}$ denote a discrete time stochastic process with the conditional mean and variance functions having parameterized by the finite dimensional vector $\theta \in \Theta \subseteq \Re^{\mathbf{m}}$ and let $\theta_{\mathbf{o}}$ denote the true value of the parameter.

Let $E[(\bullet)|\ I_{t-1}]$ or $E_{t-1}(\bullet)$ denote the mathematical expectation conditioned on the information available at time (t-1), I_{t-1}.

Definition 1. In the relationship, $u_t = Z_t\sigma_t$, the stochastic process $\{u_t(\theta),\ t \in (-\infty, \infty)\}$ follows an ARCH process if:

a. $E\ (u_t(\theta_{\mathbf{o}})\ |\ I_{t-1}) = 0$, for t = 1,2, …

b. Var $(u_t(\boldsymbol{\theta_o}) \mid I_{t-1}) = \sigma_t^2(\boldsymbol{\theta_o})$ depends non-trivially on the sigma field generated by the past observations, $\{ u_{t-1}^2(\boldsymbol{\theta_o}), u_{t-2}^2(\boldsymbol{\theta_o}), \ldots\}$.

$\sigma_t^2(\boldsymbol{\theta_o}) \equiv \sigma_t^2$ is the conditional variance of the process, conditioned on the information set I_{t-1}. The conditional variance is central to the ARCH process.

Letting $Z_t(\boldsymbol{\theta_o}) = u_t(\boldsymbol{\theta_o})/\sigma_t(\boldsymbol{\theta_o})$, t = 1,2, … we have the standardized process $\{ Z_t(\boldsymbol{\theta_o})_{t \in (-\infty, \infty)}\}$ and it follows that,

(i) $E[Z_t(\boldsymbol{\theta_o}) \mid I_{t-1}] = 0 \quad \forall t$

(ii) $\text{Var}[Z_t(\boldsymbol{\theta_o}) \mid I_{t-1}] = 1 \quad \forall t$

Thus, the conditional variance of $Z_t(.)$ is time invariant. Moreover, if we assume that the conditional distribution of $Z_t(.)$ is time invariant with a finite fourth moment, it follows from Jensen's inequality that,

$$E(u_t^4) = E(Z_t^4)E(\sigma_t^4) \geq E(Z_t^4)\,[E(\sigma_t^2)]^2 = E(Z_t^4)\,[E(u_t^2)]^2$$

with the last equality holding only when the conditional variance is constant. Assuming that $Z_t(.)$ is normally distributed, it follows that the unconditional distribution of u_t is leptokurtic.

Engle (1982) has shown that for the first order or ARCH(1) process,

$$\sigma_t^2 = \alpha_o + \alpha_1 u_{t-1}^2 \qquad (1)$$

the unconditional variance and the fourth moment for this process are, respectively, given by,

$$E(u_t^2) = \frac{\alpha_0}{1-\alpha_1}$$

$$E(u_t^4) = \left[\frac{3\alpha_0^2}{(1-\alpha_1)^2}\right]\left[\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right]$$

The condition for the variance to be finite is that $\alpha_1 < 1$ and for the fourth moment, $3\alpha_1^2 < 1$.

It implies that $E(u_t^4)/ E[(u_t^2)]^2 \geq E(Z_t^4)$. Thus, for the first order ARCH process,

$$\frac{E(u_t^4)}{(E(u_t^2))^2} = \frac{3(1-\alpha_1^2)}{(1-3\alpha_1^2)} > 3$$

This result implies that the ARCH (1) process is a heavy-tailed distribution, that is, the process generates data with fatter tails than the normal distribution. This particular characteristic of the ARCH process is relevant in modeling financial time series, like stock returns and asset prices, since these series tend to have thick-tailed distributions.

In general, the ARCH (q) process can be defined as,

$$\sigma_t^2 = \omega + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_q u_{t-q}^2 \qquad (2)$$

For this model to be well defined and have a positive conditional variance almost surely, the parameters must satisfy $\omega > 0$ and $\alpha_1, \ldots, \alpha_q \geq 0$. We will see later that for the Generalized ARCH or the GARCH process, this condition can be made less stringent.

Following the natural extension of the ARMA process as a parsimonious representation of a higher order AR process, Bollerslev (1986) extended the work of Engle to the Generalized ARCH or GARCH process. In the GARCH (p,q) process defined as,

$$\sigma_t^2 = \omega + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{q} \alpha_i \, u_{t-i}^2 \qquad (3)$$
$$\omega > 0, \alpha_i \geq 0, \beta_j \geq 0 \quad i = 1, \ldots, q \;\; j = 1, \ldots, p$$

the conditional variance is a linear function of q lags of the squares of the error terms (u_t^2) or the ARCH terms (also referred to as the “news” from the past) and p lags of the past values of the conditional variances (σ_t^2) or the GARCH terms, and a constant ω. The inequality restrictions are imposed to guarantee a positive conditional variance, almost surely.

2.2.2. The Exponential GARCH (EGARCH) Process

The GARCH process being an infinite or a higher order representation of the ARCH process captures the empirical regularities observed in the time series data such as thick-tailed distributions and volatility clustering. However, the GARCH process fails to explain the so-called “leverage effects” often observed in financial time series. The concept of leverage effects, first observed by Black (1976), refers to the tendency for changes in the stock prices to be negatively correlated with changes in the stock volatility. In other words, the effect of a shock on the volatility is asymmetric, or to put it differently, the impact of a “good news” (positive lagged residual) is different from the impact of the “bad news” (negative lagged residual).

The GARCH process, being symmetric, fails to capture this phenomenon since in the model, the conditional variance is a function only of the magnitudes of the lagged residuals and not their signs.

A model that accounts for an asymmetric response to a shock was credited to Nelson (1991) and is called an Exponential GARCH or EGARCH model. The specification for the conditional variance using the EGARCH (p,q) is,

$$\log(\sigma_t^2) = \omega + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2) + \sum_{i=1}^{q} \alpha_i \left| \frac{u_{t-i}}{\sigma_{t-i}} \right| + \sum_{k=1}^{r} \gamma_k \frac{u_{t-k}}{\sigma_{t-k}}$$

The log of the conditional variance implies that the leverage effect is exponential rather than quadratic. A commonly used model is the EGARCH (1,1) given by,

$$\log(\sigma_t^2) = \alpha_0 + \alpha_1 \left| \frac{u_{t-1}}{\sigma_{t-1}} \right| + \beta_1 \log(\sigma_{t-1}^2) + \gamma \frac{u_{t-1}}{\sigma_{t-1}} \qquad (4)$$

The presence of the leverage effects is accounted for by γ, which makes the model asymmetric. The motivation behind having an asymmetric model for volatility is that it allows the volatility to respond more quickly to falls in the prices (bad news) rather than to the corresponding increases (good news).

2.2.3. The Threshold GARCH (TARCH) Process

Another model than accounts for the asymmetric effect of the “news” is the Threshold GARCH or TARCH model due independently to Zakoïan (1994) and Glosten, Jaganathan and Runkle (1993). The TARCH (p,q) specification is given by,

$$\sigma_t^2 = \omega + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{k=1}^{r} \gamma_k u_{t-k}^2 I_{t-k}^- \qquad (5)$$

where,

$$I_{t-k}^- = \begin{cases} 1 \text{ } if \text{ } u_t < 0 \\ 0 \text{ } otherwise \end{cases}$$

In the TARCH model, “good news”, $u_{t-i} > 0$ and “bad news”, $u_{t-i} < 0$ have different effects on the conditional variance. When $\gamma_k \neq 0$, we conclude that the news impact is asymmetric and that there is a presence of leverage effects. When $\gamma_k = 0$ for all k, the TARCH model is equivalent to the GARCH model. The difference between the TARCH and the EGARCH models is that in the former the leverage effect is quadratic while in the latter, the leverage effect is exponential.

2.2.4. The Power ARCH (PARCH) Process

Most of the ARCH-type of models discussed so far deal with the conditional variance in the specification. However, when one talks of volatility the appropriate measure is the standard deviation rather than the variance as noted by Barndorff-Nielsen and Shephard (2002). A GARCH model using the standard deviation was introduced independently by Taylor (1986) and Schwert (1989). In these models, the conditional standard deviation as a measure of volatility is being modeled instead of the conditional variance. This class of models is generalized by Ding *et al*. (1993) using the Power ARCH or PARCH model. The PARCH specification is given by,

$$\sigma_t^{\delta} = \omega + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^{\delta} + \sum_{i=1}^{q} \alpha_i (|u_{t-i}| - \gamma_i u_{t-i})^{\delta} \qquad (6)$$

where,

$$\delta > 0, |\gamma_i| \leq 1 \ for \ i = 1,2,\ldots,r \ and \ \gamma_i = 0 \ for \ i > r, and \ r \leq p.$$

Note that in the PARCH model, $\gamma \neq 0$ implies asymmetric effects. The PARCH model reduces to the GARCH model when $\delta = 2$ and $\gamma_i = 0$ for all i.

III. The Realized Volatility

Let $P_{n,t}$ denote the price of an asset (say US$ 1 in Philippine Peso) at time $n \geq 0$ at day t, where n = 1,2,…,N and t=1,2,…,T. Note that when n=1, P_t is simply the inter-daily price of the asset (normally recorded as the closing price). Let $p_{n,t} = \log(P_{n,t})$, denote the natural logarithm of the price of the asset. The observed discrete time series of continuously compounded returns with N observations per day is given by,

$$r_{n,t} = p_{n,t} - p_{n-1,t} \qquad (7)$$

When n=1, we simply ignore the subscript n and $r_t = p_t - p_{t-1} = \log(P_t) - \log(P_{t-1})$ where t= 2,…,T. In this case, r_t is the time series of daily return and is also the covariance-stationary series. In (7), we assume that:

(a) $E(r_{n,t}) = 0$

(b) $E(r_{n,t}\, r_{m,s}) = 0$ for $n \neq m$ and $t \neq s$

(c) $E(r_{n,t}^2\, r_{m,s}^2) < \infty$ for n,m,s,t

Assumption (a) implies that the mean return which follows from the fact that the log prices, p_t, follow an i.i.d. random walk process without a drift,

$$p_{n,t} = p_{n-1,t} + \varepsilon_{n,t} \qquad where\ \varepsilon_{n,t} \mid I_{t-1} \sim i.i.d.(0, \sigma_t^2) \tag{8}$$

Following (8), $r_{n,t} = p_{n,t} - p_{n-1,t} = \varepsilon_{n,t}$ and thus, $E(r_{n,t}) = E(\varepsilon_{n,t}) = 0$. Assumption (b) follows from the fact that $\varepsilon_{n,t}$ are i.i.d. and from (a) which gives us $E(r_{n,t}\ r_{m,s})=E(\varepsilon_{n,t}\ \varepsilon_{m,s}) = 0$. Assumption (c) states that the variances and co-variances of the squared returns exist and are finite. This follows from the fact that $E(r_{n,t}^2\ r_{m,s}^2)=E(\varepsilon_{n,t}^2\ \varepsilon_{m,s}^2) < \infty$ for n,m,s,t.

From (7), the continuously compounded daily return (Campbell, Lo, and Mackinlay, 1997 p.11) is given by,

$$r_t = \sum_{n=1}^{N} r_{n,t} \tag{9}$$

and the continuously compounded daily squared returns is,

$$\begin{aligned} r_t^2 &= \left(\sum_{n=1}^{N} r_{n,t} \right)^2 = \sum_{n=1}^{N} r_{n,t}^2 + \sum_{n=1}^{N} \sum_{m=1}^{N} r_{n,t} r_{m,t} \\ &= \sum_{n=1}^{N} r_{n,t}^2 + 2 \sum_{n=1}^{N} \sum_{m=n+1}^{N} r_{n,t} r_{m-n,t} \end{aligned} \tag{10}$$

Note that $\sigma_t^2 = Var(r_t) = E(r_t^2)$ since $E(r_t) = 0$. From (10) and using assumption (b) of (7) above, we have,

$$\begin{aligned} &\sigma_t^2 = E(r_t^2) = E\left(s_t^2\right) \\ &where\ s_t^2 = \sum_{n=1}^{N} r_{n,t}^2 \end{aligned} \tag{11}$$

Thus, the sum of the intra-daily squared returns is an unbiased estimator of the daily population variance. The sum of the intra-daily squared returns is known as the **realized volatility** (also called the realized variance by Barndorff-Nielsen and Shephard (2002)). Given enough observations for a given trading day, the realized volatility can be computed and is a model-free estimate of the conditional variance. The properties of the realized volatility are discussed in Anderson, Bollerslev, Diebold and Labys (1999). In particular, the authors found that the realized volatility is a consistent estimator of the daily population variance, σ_t^2.

IV. The GARCH-PARK-R Model

While the concept of realized volatility does provide a highly efficient way of estimating the unknown conditional variance, the problem of generating information on the price of an asset every five minutes or so is simply enormous. An alternative measure is to use extreme values, the highest and lowest prices of an asset, to produce two intra-daily observations. The range, the difference between the highest and lowest log prices, is a good proxy for volatility. The range has the advantage of being available for researchers since high and low prices are available daily for a variety of financial time series such as price of individual stock, composite indices, Treasury bill rates, lending rates, currency prices and the like.

The log range, R_t, is defined as,

$$\begin{aligned} R_t &= \log(P_{(N),t}) - \log(P_{(1),t}) \\ &= p_{(N),t} - p_{(1),t} \qquad t = 1,2\ldots,T \qquad (12) \end{aligned}$$

where $P_{(N),t}$ is the highest price of the asset at day (or time) t and $P_{(1),t}$ is the lowest price of the asset at day t.

The log range is superior over the usual measure of volatility based on daily data, the square return $r_t^2 = \log(P_t) - \log(P_{t-1})$. Alizadeh, Brandt and Diebold (2001) noted that the log range is a better measure of volatility in the sense that the log range has fewer measurement errors compared to the squared-returns. For instance, on a given day, the price of an asset fluctuates substantially throughout the day but its closing price happens to be very close to the previous closing price. If we use the inter-daily squared return, the value will be small despite the large intra-daily price fluctuations. The log range, using the highest and lowest values, reflects a more precise price fluctuations and can indicate that the volatility for the day is high.

Moreover, the log range can be approximated by a Gaussian distribution quite well. The distribution of the range was first derived by Feller (1951) using a drift-less Brownian motion process.

As compared to the realized volatility, the log range has the advantage of being robust to certain market microstructure effects. These microstructure effects, such as the bid-ask spread, are noises that can affect the features of the time-series. The bid-ask spread is a common type of microstructure effect. Most markets require liquidity, giving way to a practice of granting monopoly rights to the so-called “market makers.” Such monopoly rights, granted by the exchange, allow the market makers to post different prices for buying and selling, they buy at a bid price P_b and sell at a higher price P_a. The difference in the prices, $P_a - P_b$ is known as the spread. Although in practice, such spread is rather small, its presence increases the volatility of the intra-daily squared returns, the input in the realized volatility, making the estimates biased upward. The log range, on the other hand, is not seriously affected by the bid-ask spread. There are other factors that create unnecessary noise in the intra-daily realized volatility such as regulatory rules imposed on the market. One such rule is the lifting of trading restrictions in the foreign exchange market for Japanese banks during the Tokyo lunch period resulting to higher volatility as documented by Anderson, Bollerslev and Das (1998).

Parkinson (1980) was the first to make use of the range in measuring volatility in the financial market. Parkinson developed the PARK daily volatility estimator based on the assumption that the intra-daily prices follow as Brownian motion. This study will make use of the PARK Range in modeling time-varying volatility. The model will be called the Generalized Auto-Regressive Conditional Heteroskedasticity Parkinson Range (GARCH-PARK-R) model.

Consider the covariance-stationary time series $\{R_{pt}\}$ where,

$$R_{P_t} = \frac{\log(P_{(N),t}) - \log(P_{(1),t})}{\sqrt{4\log(2)}} \qquad t = 1,2,\ldots,T \qquad (13)$$

R_{Pt} is the PARK-Range of the asset at time t. Moreover, let $R_{Pt} \geq 0$ for all t and that $P(R_{Pt} < \delta| R_{Pt-1}, R_{Pt-2},\ldots) > 0$ for any $\delta > 0$ and for all t. This condition states that the probability of observing zeros or near zeros in R_{Pt} is greater than zero.

Let,

$\mu_t = E[R_{Pt}| \underline{\mathbf{I}}_{t-1}]$ be the conditional mean of the PARK range

and

$\sigma_t^2 = Var[R_{Pt}|\underline{\mathbf{I}}_{t-1}]$ be the conditional variance of the PARK range

The motivation behind the GARCH-PARK-R model is the Auto-Regressive Conditional Duration (ACD) model of Engle and Russel (1998) used to model observations that arrive at irregular intervals.

Let,

$$R_{P_t} = \mu_t \varepsilon_t \qquad where \;\; \varepsilon_t \mid I_{t-1} \sim iid(1, \phi_t^2) \quad and$$

$$\mu_t = \omega + \sum_{j=1}^{q} \alpha_j R_{P_{t-j}} + \sum_{j=1}^{p} \beta_j \mu_{t-j} \tag{14}$$

The model in (14) is known as the GARCH-PARK-R process of orders p and q. The mean and variance of the PARK range are given by,

$$(a) \qquad E(R_{P_t}) = \mu_t$$

$$\begin{aligned}(b) \qquad Var(R_{P_t}) &= E(R_{P_t}^2) - [E(R_{P_t})]^2 \\ &= \mu_t^2 E\left(\varepsilon_t^2\right) - \mu_t^2 \\ &= \mu_t^2 (\phi_t^2 + 1) - \mu_t^2 = \mu_t^2 \phi_t^2 \end{aligned} \tag{15}$$

The GARCH-PARK-R model is similar to the Conditional Auto-Regressive Range (CARR) model suggested by Chou (2003). Two differences between this study and that of Chou's are to be clarified. First, this study uses the Parkinson range to study volatility instead of the usual log range (Chou's measure). The Parkinson range has been found to be a better estimator of volatility (standard deviation). The second difference is the use of the data, this study makes use of the daily data while Chou's paper used weekly data. Weekly data may have distorted estimates due to the presence of aggregation effect that is why the author of this paper used the daily data instead.

For the density function of ε_t in (14), this study follows the suggestion of Engle and Russel (1998) and Engle and Gallo (2003) of using the gamma density,

$$f(\varepsilon_t \mid I_{t-1}) = \frac{\left(1/\beta\right)^{\alpha}}{\Gamma(\alpha)} \varepsilon_t^{\alpha-1} \exp\left\{-\frac{\varepsilon_t}{\beta}\right\} \tag{16}$$

Since E(ε_t) = $\alpha\beta$ = 1 (by assumption in (14)), it implies that $\alpha = 1/\beta$. Thus (16) now becomes,

$$
\begin{aligned}
f(\varepsilon_t \mid I_{t-1}) &= \frac{\alpha^{\alpha}}{\Gamma(\alpha)} \varepsilon_t^{\alpha-1} \exp\{-\alpha\varepsilon_t\} \\
\Rightarrow f(R_{P_t} \mid I_{t-1}) &= f\left(\frac{R_{P_t}}{\mu_t} \mid I_{t-1}\right) = \frac{1}{\Gamma(\alpha)} \alpha^{\alpha} \left(\frac{R_{P_t}}{\mu_t}\right)^{\alpha-1} \exp\left\{-\alpha\left(\frac{R_{P_t}}{\mu_t}\right)\right\} \\
&= \frac{1}{\Gamma(\alpha)} \alpha^{\alpha} \left(R_{P_t}\right)^{\alpha-1} (\mu_t)^{-\alpha} \exp\left\{-\alpha\left(\frac{R_{P_t}}{\mu_t}\right)\right\} \\
&= \frac{\left(\alpha/\mu_t\right)^{\alpha}}{\Gamma(\alpha)} \left(R_{P_t}\right)^{\alpha-1} \exp\left\{-R_{P_t}\left(\frac{\alpha}{\mu_t}\right)\right\} \qquad (17)
\end{aligned}
$$

From (17), the conditional mean and variances of R_{Pt} are,

$$
\begin{aligned}
E\left(R_{P_t} \mid I_{t-1}\right) &= \frac{\alpha}{\alpha/\mu_t} = \mu_t \\
Var(R_{P_t} \mid I_{t-1}) = \frac{\alpha}{\left(\alpha/\mu_t\right)^2} &= \frac{(\mu_t)^2}{\alpha}
\end{aligned}
$$

The density function in (17) approaches the Gaussian density as α increases. Moreover, the likelihood function is given by,

$$
L = \prod_{t=1}^{T}\left[\frac{1}{\Gamma(\alpha)} \alpha^{\alpha} \left(R_{P_t}\right)^{\alpha-1} (\mu_t)^{-\alpha} \exp\left\{-\alpha\left(\frac{R_{P_t}}{\mu_t}\right)\right\}\right] \qquad (18)
$$

If the parameter of interest are only those that define μ_t in (14), denoted by $\mu_t(\underline{\theta})$, then the log likelihood can be simplified into,

$$\log\, L \quad = \gamma - \alpha \sum_{t=1}^{T} \log(\mu_t) - \alpha \sum_{t=1}^{T} \left(\frac{R_{P_t}}{\mu_t} \right)$$

$$= \gamma - \alpha \sum_{t=1}^{T} \left[\log(\mu_t(\underline{\theta})) + \frac{R_{P_t}}{\mu_t(\underline{\theta})} \right] \qquad (19)$$

$where\ \gamma = \gamma(\alpha, R_{P_t})$

Taking the derivative of the log likelihood function with respect to $\underline{\theta}$, we have,

$$\frac{\partial \log(L)}{\partial \underline{\theta}} = -\alpha \sum_{t=1}^{T} \left[\frac{1}{\mu_t(\underline{\theta})} \frac{\partial \mu_t(\underline{\theta})}{\partial \underline{\theta}} - \frac{R_{P_t}}{\mu_t^2(\underline{\theta})} \frac{\partial \mu_t(\underline{\theta})}{\partial \underline{\theta}} \right] = 0$$

$$\Rightarrow$$

$$\sum_{t=1}^{T} \left[\frac{R_{P_t}}{\mu_t^2(\underline{\theta})} - \frac{1}{\mu_t(\underline{\theta})} \right] \frac{\partial \mu_t(\underline{\theta})}{\partial \underline{\theta}} = 0$$

$$\Rightarrow$$

$$\sum_{t=1}^{T} \left[\frac{R_{P_t} - \mu_t(\underline{\theta})}{\mu_t^2(\underline{\theta})} \right] \frac{\partial \mu_t(\underline{\theta})}{\partial \underline{\theta}} = 0 \qquad (20)$$

The parameter vector $\underline{\theta}$ in (20) can be estimated numerically using some iterative algorithms such as the Marquardt or the BHHH.

An easier way of estimating the parameter vector $\underline{\theta}$ is to apply the method of estimating the parameters of a GARCH (p,q) process. Recall that in the GARCH (p,q) process discussed in section 2,

$$u_t = \sigma_t Z_t \qquad Z_t \sim N(0,1)$$

$$and$$

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

For the GARCH-PARK-R process let,

$$
\begin{aligned}
&\sqrt{R_{P_t}} = \sqrt{\mu_t}\,\nu_t \qquad\qquad \nu_t \sim iid(0,1)\\
&E\left(\sqrt{R_{P_t}} \mid I_{t-1}\right) = \sqrt{\mu_t}\,E(\nu_t) = 0\\
&and \qquad Var\left(\sqrt{R_{P_t}} \mid I_{t-1}\right) = \mu_t E\left(\nu_t^2\right) = \mu_t\\
&with\\
&\mu_t = \omega + \sum_{j=1}^{q} \alpha_j R_{P_{t-j}} + \sum_{j=1}^{p} \beta_j \mu_{t-j} \qquad given\ in\ (14)
\end{aligned}
$$

Thus, an analogous method of estimating the parameter vector $\underline{\theta}$ is to estimate the variance equation for the positive square root of the PARK R using GARCH (p,q) specification with zero in the mean specification. The Quasi-Maximum Likelihood estimators are consistent and distributed as Gaussian asymptotically even if the probability density function of the error is mis-specified following the results of Lee and Hansen and Lumsdaine for the GARCH (1,1) and Berkes *et al* for the GARCH (p,q) process. Obviously, if the correct specification is satisfied, for instance using the gamma distribution, the QMLE is the MLE and the estimators are asymptotically efficient. Therefore, a trade-off has to be made. This study will make use of the QMLE and aspire for consistency and asymptotic normality of the estimators. The data and the results of the empirical analysis are discussed in the next chapter.

V. Empirical Results

This chapter discusses the results of forecasting the conditional variance using the different ARCH and GARCH-PARK-R models. In this study, a total of 77 models were estimated: 68 ARCH-type models and 9 GARCH-PARK-R models. The model specifications are provided in Tables 1A and 1B below.

Table 1A. Specification for ARCH-type Models *

Model	Specification	Model	Specification
1	ARCH (1)	10	TARCH (1,1)
2	GARCH (1,1)	11	TARCH (1,2)
3	GARCH (1,2)	12	TARCH (2,1)
4	GARCH (2,1)	13	TARCH (2,2)
5	GARCH (2,2)	14	PARCH (1,1)
6	EGARCH (1,1)	15	PARCH (1,2)
7	EGARCH (1,2)	16	PARCH (2,1)
8	EGARCH (2,1)	17	PARCH (2,2)
9	EGARCH (2,2)		

* The 17 models are estimated via the MLE using the Gaussian, Student's t and the Generalized Error Distribution and using the QMLE resulting to 68 models.

Table 1B. Specification for GARCH PARK R Models*

Model	Specification	Model	Specification
1	ARCH (1)	6	EGARCH (1,1)
2	GARCH (1,1)	7	EGARCH (1,2)
3	GARCH (1,2)	8	EGARCH (2,1)
4	GARCH (2,1)	9	EGARCH (2,2)
5	GARCH (2,2)		

* The models are estimated using QMLE.

These models were estimated to fit the daily returns of the peso-dollar exchange rate from January 02, 1997 to December 05, 2003, a total of 1730 observations.

Following the approach of Hansen and Lunde (2001), the time series was divided into two sets, an estimation period and an evaluation period.

$$\mathbf{t} = \underbrace{-\mathbf{T}+\mathbf{1},\ldots,\mathbf{0}}_{\text{estimation period}} \qquad \underbrace{\mathbf{1},\mathbf{2},\ldots,\mathbf{n}}_{\text{evaluation period}}$$

The parameters of the volatility models are estimated using the first T inter-daily observations and the estimates of the parameters are used to make forecasts of for the remaining n periods. The estimation period made use of daily returns from January 02, 1997 to December 27, 2002, a total of 1493 observations.

In the evaluation period the daily volatility is estimated using the square of the Parkinson R, defined in (31). The square of the PARK R serves as the proxy for the unknown conditional variance. The evaluation period makes use of daily returns from January 02, 2003 to December 05, 2003, a total of 237 observations.

5.2. Loss Functions

Let h denote the number of competing forecasting models. The j^th model provides a sequence of forecasts for the conditional variance,

$$\hat{\sigma}^2_{j,1}, \hat{\sigma}^2_{j,2}, \ldots, \hat{\sigma}^2_{j,n} \qquad j = 1,2,\ldots,h$$

that will be compared to the square of the Parkinson range, the proxy of the intra-daily calculated volatility,

$$R^2_{P_1}, \ldots, R^2_{P\,n}$$

The forecast of j^th model leads to the observed loss,

$$L_{j,t}(\hat{\sigma}^2_{j,t}, R^2_{P_t}) \quad j = 1,2,\ldots,77 \quad and \quad t = 1,2,\ldots,237$$

In this study, five (5) different loss functions are used to evaluate the forecasting performance of the different models. The loss functions are:

$$MAD_1 = \frac{\sum_{t=1}^{n} \left| R_{P_t} - \hat{\sigma}_t \right|}{n} \qquad (21)$$

$$MAD_2 = \frac{\sum_{t=1}^{n} \left| R_{P_t}^2 - \hat{\sigma}_t^2 \right|}{n} \qquad (22)$$

$$MSE_1 = \frac{\sum_{t=1}^{n} \left(R_{P_t} - \hat{\sigma}_t \right)^2}{n} \qquad (23)$$

$$MSE_2 = \frac{\sum_{t=1}^{n} \left(R_{P_t}^2 - \hat{\sigma}_t^2 \right)^2}{n} \qquad (24)$$

$$R^2 LOG = \frac{\sum_{t=1}^{n} \left[\log\left(R_{P_t}^2 / \hat{\sigma}_t^2 \right) \right]^2}{n} \qquad (25)$$

The criteria (21) to (24) are the usual mean absolute deviations and mean square errors using the forecasts of the conditional standard deviation and the conditional variance.

Criterion (25) is equivalent to the R^2 criterion using the regression equation,

$$\log(R_{P_t}^2) = a + b \log(\hat{\sigma}_t^2) + \varepsilon_t \quad t = 1,2,\ldots,237$$

discussed in Engle and Patton (2001) and Taylor (1999).

The results of the forecasting performance are provided in Table 2 of the Appendix of this study. The “best” 10 models are shown in Table 3 below. The best over-all ARCH model is the TARCH (2,2) model with the Student’s t as the underlying distribution. The second “best” model is the PARCH (2,2) model, also using the Student’s t distribution.

Table 3. Forecasting Performance of the Top 10 ARCH Models

MAD1		MAD2		MSE1		MSE2		R2LOG	
Model	**Mean**	**Model**	**Mean**	**Model**	**Mean**	**Model**	**Mean**	**Model**	**Mean**
t 13	9.1000E-04	t 13	5.3400E-06	t 13	1.7100E-06	t 13	9.1500E-11	t 13	1.1105E+00
t 17	9.5000E-04	t 17	5.7200E-06	t 17	1.9000E-06	GED 7	9.6700E-11	t 17	1.1435E+00
GED 12	1.0320E-03	GED 12	6.2700E-06	GED 12	2.1200E-06	t 17	1.0900E-10	GED 15	1.3438E+00
GED 4	1.0360E-03	GED 4	6.3900E-06	GED 4	2.1800E-06	GED 12	1.1200E-10	GED 4	1.3476E+00
GED 15	1.0460E-03	GED 15	6.5400E-06	GED 5	2.2300E-06	GED 4	1.1500E-10	GED 5	1.3494E+00
GED 10	1.0480E-03	GED 5	6.5700E-06	GED 7	2.2300E-06	GED 5	1.2200E-10	GED 10	1.3549E+00
GED 11	1.0570E-03	GED 19	6.6100E-06	GED 15	2.2600E-06	GED 15	1.2600E-10	GED 12	1.3561E+00
GED 5	1.0580E-03	GED 11	6.6500E-06	GED 10	2.3100E-06	GED 6	1.2800E-10	GED 11	1.3590E+00
GED 14	1.0580E-03	GED 14	6.6800E-06	GED 11	2.3200E-06	GED 17	1.3000E-10	GED 14	1.3647E+00
GED 3	1.0800E-03	GED 3	6.8200E-06	GED 14	2.3200E-06	GED 10	1.3100E-10	GED 3	1.3876E+00

From Table 3, it is interesting to note that models using the Generalized Error Distribution performed relatively well using the five forecasting criteria, with 8 out of 17 models landing in the top 10 models. In general, the models with relatively superior forecasting performance, using the peso-dollar exchange rate, are those that accommodate the leverage effects such as the TARCH, PARCH and EGARCH. However, while the correct specification of the volatility is important, one must also consider the distribution used in estimating the parameters of the model.

The results in Table 2 showed that volatility models that assumed the Gaussian distribution or those that used the QMLE performed worst compared to models that assumed the Student’s t or Generalized Error distributions. Therefore, it is important to correctly specify the entire distribution and not only to focus on the specification of the volatility, even if it is the object of interest. A similar observation was made in the study of Hansen and Lunde (2001).

Using the five criteria discussed above the forecasting performance of the GARCH – PARK – R models are given in Table 4. The top three models are GARCH (1,2), (2,1) and (1,1). It should be noted that while the GARCH (1,2) and the GARCH (2,1) have outperformed, albeit slightly, the GARCH (1,1), the latter is preferred since the coefficients α and β are significantly different from zero.

Table 4. Forecasting Performance of the GARCH-PARK-R Models

PARK R Model	MAD1	MAD2	MSE1	MSE2	R2LOG
	Mean	Mean	Mean	Mean	Mean
GARCH (1,2)	9.2600E-04	5.2700E-06	1.4200E-06	7.3700E-11	1.0403E+00
GARCH (2,1)	9.3800E-04	5.3700E-06	1.4600E-06	7.6400E-11	1.0501E+00
GARCH (1,1)	9.4300E-04	5.4100E-06	1.4800E-06	7.7900E-11	1.0540E+00
GARCH (2,2)	9.4400E-04	5.4200E-06	1.4800E-06	7.8200E-11	1.0544E+00
EGARCH (1,1)	1.0600E-03	6.1600E-06	1.7400E-06	8.6600E-11	1.2204E+00
EGARCH (2,1)	1.0610E-03	6.1800E-06	1.7500E-06	8.7800E-11	1.2221E+00
EGARCH (1,2)	1.0620E-03	6.1900E-06	1.7600E-06	8.8800E-11	1.2231E+00
ARCH (1)	1.1760E-03	7.0300E-06	2.1300E-06	1.1300E-10	1.3375E+00
EGARCH (2,2)	1.2450E-03	7.5800E-06	2.4900E-06	1.3000E-10	1.4849E+00

As expected, the GARCH-PARK-R models performed better than most of the ARCH-type models. This is expected since the proxy for the conditional variance in the evaluation period is the square of the Parkinson range. However, it is interesting to note that the forecasting performance of the “best” ARCH-type model, the TARCH (2,2) model with a student’s t distribution, is relatively near the “best” GARCH-PARK-R model. The results somewhat provide an assurance that volatility models using inter-daily data can forecast the conditional variance rather well (at least using the Parkinson range).

VI. Conclusion

This paper introduced a relatively simple, yet efficient, model to describe the variation in volatility of the peso-dollar exchange rate using intra-daily returns. The Generalized Auto-Regressive Conditional Heteroskedasticity Parkinson Range (GARCH-PARK-R) model can actually produce volatility estimates that are relatively superior than the ARCH class of models using inter-daily returns. The GARCH-PARK-R model is a good alternative to the so-called Realized Volatility that makes use of large quantity of intra-daily data, something that is difficult to obtain in emerging markets such as the Philippines.

REFERENCES

Andersen T.G. and Bollerslev, T. (1998), "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts", *International Economic Review,* 39, 885-905.

Anderson, T., Bollerslev, T., Diebold, F. X., and Labys, P., "*The Distribution of Exchange Rate Volatility*", working paper 6961, National Bureau of Economic Research, February 1999.

Barndorff-Nielsen, O. E. and Shephard, N. (2002), "Estimating Quadratic Variation using Realized Variance", *Journal of Applied Econometrics*, 17, 457-477.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 307-327.

Bollerslev, T. and Wooldridge, J. M. (1992), "Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances", *Economic Reviews*, 11, 143-172.

Busch, T., "*Finite Sample Properties of GARCH Quasi-Maximum Likelihood Estimators and Related Test Statistics"*, working paper, University of Aarhus, October 2003.

Campbell, J. Y., Lo, A. W., and Mackinlay, A. C. (1997), *The Econometrics of Financial Markets.* USA: Princeton University Press.

Chou, R. (2003), *“Forecasting Financial Volatilities with Extreme Values: The Conditional Autoregressive Range (CARR) Model*”, paper, Institute of Economics, Academia Sinica.

Ding, Z., Engle, R.F., and Granger, C. W. J. (1993) “Long Memory Properties of Stock Market Returns and a New Model”, *Journal of Empirical Finance*, 1, 83-106.

Engle, R. F. (1982), “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation”, *Econometrica*, 50, 987-1007.

Engle, R. F. (2002), “New Frontiers for ARCH Models”, *Journal of Applied Econometrics*, 17, 425-446.

Engle, R. F. and Gallo, G. (2003), “*A Multiple Indicators Model for Volatility Using Intra-Daily Data”,* working paper 10117, National Bureau of Economic Research, November 2003.

Engle, R. F. and Patton, A. J. (2001), *“What Good is a Volatility Model?”* unpublished manuscript, Department of Finance, Stern School of Business, New York University.

Engle, R. F. and Russell, J. (1998), “Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data”, *Econometrica*, 66, 1127-1162.

Glosten, L. R., Jagannathan, R., and Runkle, D. (1993) “On the relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks”, *Journal of Finance*, 48, 1779-1801.

Hansen, P. and Lunde, A. (2001), *“A Forecast Comparison of Volatility Models: Does anything Beat a GARCH (1,1)?”*, working paper, Department of Economics, Brown University, November 2001.

Lee, S. and Hansen B. E. (1994), “Asymptotic Theory for the GARCH (1,1) Quasi-Maximum Likelihood Estimator”, *Economic Theory*, 10, 29-52.

Lumsdaine, R. L. (1996), “Consistency and Asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH (1,1) and Covariance Stationarity GARCH (1,1) Models”, *Econometrica*, 64, 575, 596.

Nelson, D. (1991), “Conditional Heteroskedasticity in Asset Returns: A New Approach”, *Econometrica*, 59, 347-370.

Parkinson, M. (1980), “The Extreme Value Method for Estimating the Variance of the Rate of Return”, *Journal of Business*, 53, 61-65.

Taylor, J. W. (1999), “Evaluating Volatility and Interval Forecasts”, *Journal of Forecasting*, 18, 11-128.

Taylor, S.J. (1986), *Modeling Financial Time Series*. New York: John Wiley and Sons.