

# A Forecast Comparison of Financial Volatility Models: GARCH (1,1) is not Enough<sup>1</sup>

Dennis S. Mapa<sup>2</sup>

## ABSTRACT

Asset allocation and risk calculations depend largely on volatility models. The parameters of the volatility models are estimated using either the Maximum Likelihood (ML) or the Quasi-Maximum Likelihood (QML). By comparing the out-of-sample forecasting performance of 68 ARCH-type models using inter-daily data on the peso-dollar exchange rate, this study shows that it is important to correctly specify the distribution of the asset returns and not only focus on the specification of the volatility. The forecasts are compared to the Parkinson Range, an alternative to the Realized Volatility.

**KEYWORDS:** Volatility, ARCH, Parkinson Range

## 1. INTRODUCTION

Since the introduction of the seminal paper on the ARCH model by Robert Engle (1982), various works on the financial and time series econometrics have been dominated by the extensions of the ARCH process. This area of research has been growing very fast over the years and while one might think that the frontier of this research program has already been reached and topics already exhausted, new interesting papers related to the subject are still being published in rapid succession.

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 the good models from the bad ones. Anderson and Bollerslev (1998) introduced the concept of “realized volatility” from which evaluation of the ARCH volatility models are to be made. Realized volatility models are calculated from high-frequency intra-daily data, rather than inter-daily data. 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.

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 prices, is a good proxy for volatility. The range has the advantage of being available to researchers since high and low prices are available daily for a variety of financial time series such as price of

---

<sup>1</sup> This paper was derived mainly from the empirical results in Chapter 5 of the M.S. Thesis entitled “The Generalized AutoRegressive Conditional Heteroskedasticity Parkinson Range(GARCH-PARK-R) Model for Forecasting Financial Volatility.” This paper which won the Best paper for the Student Session category is reprinted with permission from the 9<sup>th</sup> National Convention on Statistics Publication – Convention Papers (Volume II).

<sup>2</sup> Assistant Professor, School of Statistics, University of the Philippines, Diliman, 1101, Quezon City, Philippines, Phone: (632) 9280881, Email: [csmapa@up.edu.ph](mailto:csmapa@up.edu.ph)

individual stock, composite indices, Treasury bill rates, lending rates, currency prices and the like. This paper proposes the use of the Range, in particular the Parkinson Range (Parkinson, 1980), as a benchmark from which to compare forecasts of the different volatility models. The range as a proxy for the standard deviation is rather popular in statistics, especially in the area of quality control. The advantage of using the range is that we only need to record the extreme values of the data set – the lowest and highest – and these values are readily available for most financial time series.

This paper is organized as follows: section 1 serves as introduction, section 2 introduces the concept of Realized Volatility and discusses the Parkinson Range. Section 3 provides the empirical discussion and section 4 concludes.

### 1.1. The AutoRegressive Conditional Heteroskedasticity (ARCH) Process

Let  $\{u_t(\theta), t \in (\dots, -1, 0, 1, \dots)\}$  denote a discrete time stochastic process and let  $E[(\bullet) | I_{t-1}]$  or  $E_{t-1}(\bullet)$  denote the mathematical expectation conditioned on the information available at time  $(t-1)$ ,  $I_{t-1}$ . 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.  $E(u_t(\theta_0) | I_{t-1}) = 0$ , for  $t = 1, 2, \dots$
- b.  $\text{Var}(u_t(\theta_0) | I_{t-1}) = \sigma_t^2(\theta_0)$  depends non-trivially on the sigma field generated by the past observations,  $\{u_{t-1}^2(\theta_0), u_{t-2}^2(\theta_0), \dots\}$ .

$\sigma_t^2(\theta_0) \equiv \sigma_t^2$  is the conditional variance of the process, conditioned on the information set  $I_{t-1}$ . The conditional variance is central to the ARCH process. The ARCH (q) process can be defined as,

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

For this model to be well defined and have a positive conditional variance almost surely, the parameters must satisfy  $\omega > 0$  and  $\alpha_1, \dots, \alpha_q \geq 0$ .

### 1.2. Extensions of the ARCH Process

This section discusses some of the common extensions of the ARCH process originally proposed by Robert F. Engle. The motivations behind these new models range from finding a more parsimonious model like the Generalized ARCH or GARCH process to explaining some of the stylized facts in the financial time series that are not captured by the original ARCH process, like the Exponential GARCH (EGARCH) process.

### 1.2.1. The Generalized ARCH (GARCH) Process

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,

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

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

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 ( $\sigma_t^2$ ) or the GARCH terms, and a constant  $\omega$ . The inequality restrictions are imposed to guarantee a positive conditional variance, almost surely.

Often, the GARCH (1,1) process,  $\sigma_t^2 = \omega + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2$ , is sufficient enough to explain the characteristics of the time series and is a popular model in econometrics and financial time series (Hansen and Lunde, 2001).

### 1.2.2. The Exponential GARCH (EGARCH) Process

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). 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. A commonly used model is the EGARCH (1,1) given by,

$$\log(\sigma_t^2) = \alpha_0 + \alpha_1 \left| \frac{u_{t-1}}{\sigma_{t-1}} \right| + \beta_1 \log(\sigma_{t-1}^2) + \gamma \frac{u_{t-1}}{\sigma_{t-1}} \quad (3)$$

The presence of the leverage effects is accounted for by  $\gamma$ , 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).

### 1.2.3. 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 Zakoian (1994) and Glosten, Jagannathan and Runkle (1993). The TARCH (p,q) specification is given by,

$$\sigma_t^2 = \omega + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{k=1}^r \gamma_k u_{t-k}^2 I_{t-k}^- \quad (4)$$

where,

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

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.

### 1.2.4. 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). 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,

$$\sigma_t^\delta = \omega + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta + \sum_{i=1}^q \alpha_i (|u_{t-i}| - \gamma_i u_{t-i})^\delta \quad (5)$$

where,

$$\delta > 0, |\gamma_i| \leq 1 \text{ for } i = 1, 2, \dots, r \text{ and } \gamma_i = 0 \text{ for } i > r, \text{ and } r \leq p.$$

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  $i$ .

The parameters of the ARCH models are estimated from the disturbance term  $u_t$  of the mean specification via the maximum likelihood estimation (MLE) procedure. The commonly used density functions are the Gaussian, Student's  $t$  and the Generalized Error distributions. The last two distributions are popular because both have fatter tails than the Gaussian distribution. The derivation of the MLE is omitted in this paper but can be made available upon request from the author.

An alternative type of estimation procedure is known as the Quasi-Maximum Likelihood Estimation (QMLE). The idea behind the QMLE is that even if the true probability density function family is misspecified, it is possible for an extremum estimator based on the likelihood function associated with the misspecified probability density function to possess good asymptotic properties such as consistency and asymptotic normality. The assumption in the QMLE is to correctly specify the mean and variance of the random variable  $Z_t$  in the ARCH process (page 2) and use the Gaussian log likelihood function as a vehicle to estimate the parameters. Bollerslev and Wooldridge (1992) first derived the QMLE for a wide range of the ARCH models. Lee and Hansen (1994) and Lumsdaine (1996) derived the consistency property of the estimators of the GARCH (1,1) process. Berkes, Horvath and Kokoszka (2003) extended the work of Lee and Hansen and Lumsdaine to the case of the GARCH (p,q) process.

## 2. REALIZED VOLATILITY

Difficulty in evaluating and comparing volatility models is due to the fact that volatility is not directly observable. Since there is no "benchmark" from which we can compare the forecasts of the different volatility models, identifying "bad" models from good ones is quite difficult. Anderson and Bollerslev (1998) introduced the concept of "realized volatility" from which evaluation of the ARCH volatility models are to be made. Realized 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 to 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.

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  $n = 1, 2, \dots, N$  and  $t = 1, 2, \dots, T$ . 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,

$$r_{n,t} = p_{n,t} - p_{n-1,t} \quad (6)$$

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  $t = 2, \dots, T$ . In this case,  $r_t$  is the time series of daily return and is also the covariance-stationary series. From (6), the continuously compounded daily return (Campbell, Lo, and Mackinlay, 1997 p.11) is given by,

$$r_t = \sum_{n=1}^N r_{n,t} \quad (7)$$

and the continuously compounded daily squared returns is,

$$r_t^2 = \left( \sum_{n=1}^N r_{n,t} \right)^2 = \sum_{n=1}^N r_{n,t}^2 + \sum_{n=1}^N \sum_{m=1}^N r_{n,t} r_{m,t} = \sum_{n=1}^N r_{n,t}^2 + 2 \sum_{n=1}^N \sum_{m=n+1}^N r_{n,t} r_{m-n,t} \quad (8)$$

It can be shown that,

$$\sigma_t^2 = E(r_t^2) = E(s_t^2) \quad (9)$$

$$\text{where } s_t^2 = \sum_{n=1}^N r_{n,t}^2$$

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 particular 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 (2001). In particular, the authors found that the realized volatility is a consistent estimator of the daily population variance,  $\sigma_t^2$ . 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 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.

Parkinson (1980) was the first to make use of the range in measuring volatility in the financial market. Parkinson developed the Parkinson Range (PARK-R) daily volatility estimator based on the assumption that the intra-daily prices follow as Brownian motion. As compared to the realized volatility, the 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 range, on the other hand, is not seriously affected by the bid-ask spread. Consider the covariance-stationary time series  $\{R_{pt}\}$  where,

$$R_{P_t} = \frac{\log(P_{(N),t}) - \log(P_{(1),t})}{\sqrt{4 \log(2)}} \quad t = 1, 2, \dots, T \quad (10)$$

$R_{P_t}$  is the PARK-R of the asset at time t.

### 3. DATA ANALYSIS

#### 3.1. Model Specifications

This section discusses the results of forecasting the conditional variance using the different ARCH models. In this study 68 ARCH-type models were estimated. The model specifications are provided in Tables 1 below. Each specification is estimated using the Maximum Likelihood, with Gaussian, Student's t and Generalized Error distributions and the Quasi-Maximum Likelihood. 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.

**Table 1.** 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.

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

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

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 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 Range. 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.

### 3.2. Measures to Evaluate the Forecasting Performance

The main objective of building volatility models is to forecast future volatility. Given a number of competing models, there is a need to evaluate the forecasting performance of the models to segregate “good” models from “bad” ones. This section discusses some of the commonly used measures to evaluate the forecasting performance of the volatility models. Let h denote the number of competing forecasting models. The j<sup>th</sup> model provides a sequence of forecasts for the conditional variance,

$$\hat{\sigma}_{j,1}^2, \hat{\sigma}_{j,2}^2, \dots, \hat{\sigma}_{j,n}^2 \quad j = 1, 2, \dots, h$$

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

$$R_{P1}^2, \dots, R_{Pn}^2$$

The forecast of  $j^{\text{th}}$  model leads to the observed loss,

$$L_{j,t}(\hat{\sigma}_{j,t}^2, R_{p_t}^2) \quad j = 1, 2, \dots, 68 \quad \text{and} \quad t = 1, 2, \dots, 237$$

In this study, five (5) different loss functions are used to evaluate the forecasting performance of the different models. The loss functions are based on the mean absolute deviations using the estimated conditional standard deviation (MAD1) and variance (MAD2), mean square error based on the conditional standard deviation (MSE1) and variance (MSE2) and a criterion equivalent to the  $R^2$  criterion using the regression equation,

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

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

### 3.3. Empirical Results

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. 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 of the empirical analysis 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).

## IV. CONCLUSIONS

This study compared a large number of ARCH-class of volatility models using inter-daily returns of the peso-dollar exchange rate. The estimated models are compared in terms of their out-of-sample forecasting performance to characterize the variation in the volatility. The Parkinson Range is used as the estimate of the daily volatility where comparison of the different volatility models was made. The empirical analysis showed that it is important to correctly specify the entire distribution of the volatility model and not only focus on the specification of the volatility.

It is gratifying to note that, 22 years after the original ARCH paper of Engle and a year after Engle won the Nobel Prize, doing research in the area of volatility estimation is still both dynamic and challenging.



## ACKNOWLEDGEMENT

The author wishes 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 author's responsibility.

## 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. (2001), "Modeling and Forecasting Realized Volatility", working paper 8160, National Bureau of Economic Research, March 2001.
- BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2002), "Estimating Quadratic Variation using Realized Variance", *Journal of Applied Econometrics*, 17, 457-477.
- BLACK, F. (1976), "Studies of Stock Price Volatility Changes", *Proceedings from the American Statistical Association, Business and Economic Section*, 177-181.
- 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.
- CAMPBELL, J. Y., LO, A. W., and MACKINLAY, A. C. (1997), *The Econometrics of Financial Markets*. USA: Princeton University Press.
- 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. and PATTON, A. J. (2001), "What Good is a Volatility Model?" unpublished manuscript, Department of Finance, Stern School of Business, New York University.
- 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), *Modelling Financial Time Series*. New York: John Wiley and Sons.