

HERCMA 2001
E. A. Lipitakis (Editor)
2001, L.E.A.
Printed in Hellas

MEASURING THE PREDICTIVE ACCURACY OF THE CORRELATED GAMMA RATIO METHOD FOR MODEL SELECTION

Stavros Degiannakis* and Evdokia Xekalaki*

Abstract—Autoregressive Conditional Heteroscedasticity (ARCH) models have successfully been applied in order to predict asset return volatility. Predicting volatility is of great importance in pricing financial derivatives, selecting portfolios, measuring and managing investment risk more accurately. In this paper, a number of ARCH models are examined in the framework of a method for model selection based on the Correlated Gamma Ratio (CGR) distribution and their ability to predict future volatility is examined. According to this method, the ARCH model with the lowest sum of squared standardized forecasting errors is selected for predicting future volatility. A number of evaluation criteria are used to examine the performance of a model to predict future volatility, for forecasting horizons ranging from one day to one hundred days ahead. The results show that the CGR model selection procedure has a satisfactory performance in selecting that model that generates “better” volatility predictions. It appears, therefore, that it can be regarded as a tool in guiding one’s choice of the appropriate model for predicting future volatility, with applications in evaluating portfolios, managing financial risk and creating speculative strategies with options.

Index terms—ARCH models, forecast volatility, model selection, predictability, correlated gamma ratio distribution.

I. Introduction

To evaluate their accuracy, volatility forecasts have to be compared with realized volatility, which cannot be observed. In the literature, it is common practice to refer the observed squared returns as the actual volatility. In this paper, a number of evaluation criteria are used to examine the ability of the CGR model selection method to indicate that ARCH model that generates “better” volatility predictions, for a forecasting horizon ranging from one day to one hundred days ahead. In the sequel, it is shown that the CGR method has a satisfactory performance in selecting that ARCH model that tracks realized volatility closer, for a forecasting horizon ranging from 16 days to 36 days ahead. So, it is possible to use this model selection method in financial applications requiring volatility forecasts for a

period longer than one day, i.e. option pricing, risk management. In section II of the paper, the ARCH process is presented. Section III describes the CGR model selection method in the context of ARCH models. Section IV provides a brief description of the evaluation criteria and the realized volatility measures considered. In section V the ability of the method proposed to select the ARCH model that generates “better” predictions of the volatility, is examined. The results are based on Degiannakis and Xekalaki [13]. Finally, in section VI a brief discussion of the results is provided.

II. The Autoregressive Conditional Heteroscedasticity (ARCH) Process

For P_t denoting the price of an asset at time t , let $y_t = \ln(P_t/P_{t-1})$ denote the continuously compounded return series of interest. The return series is decomposed into two parts, the predictable and unpredictable component:

$$y_t = E(y_{t|t-1}) + \varepsilon_t, \quad (1)$$

where $E(y_{t|t-1})$ is the conditional mean of return at period t depending upon the information set available at time $t-1$ and ε_t is the prediction error. Usually, the predictable component is either the overall mean or a first order autocorrelated process (imposed by non-synchronous trading¹). The conditional mean, unfortunately, does not have the ability to give useful predictions. That is why modern financial theory assumes the asset returns are unpredictable. Before the start of the 1980’s, the view taken about returns in financial markets was that they behave as random walks and the Brock et al. [9] (BDS) statistic has widely been used to test the null hypothesis that asset

¹ According to Campbell et al. [10], “The non-synchronous trading or non-trading effect arises when time series, usually asset prices, are taken to be recorded at time intervals of one length when in fact they are recorded at time intervals of other, possible irregular lengths.” For more details on non-synchronous trading see Scholtes and Williams [34], Dimson [14], Cohen et al. [11], Lo and MacKinlay [26], [27], Campbell et al. [10].

* Athens University of Economics and Business

returns are independently and identically distributed. This hypothesis, however, has been rejected in a vast number of applications. A rejection of the null hypothesis is consistent with some types of dependence in the data, which could result in from a linear stochastic system, a nonlinear stochastic system, or a nonlinear deterministic system. Thus, a question arises: "Are the nonlinearities connected with the conditional mean (so, as to be used to predict future returns) or with higher order conditional moments?" Artificial neural networks², chaotic dynamical systems³, nonlinear parametric and nonparametric models⁴ are some examples from the literature dealing with conditional mean predictions. ARCH models⁵ and Stochastic Volatility models⁶ are examples from the literature dealing with conditional variance modeling. However, no nonlinear models that can significantly outperform even the simplest linear model in out-of-sample forecasting seem to exist in the literature (neither in the field of stochastic nonlinear models nor in the field of deterministic chaotic systems). On the other hand, the ARCH processes and Stochastic Volatility models appear to be more appropriate to interpret nonlinearities in financial systems on the basis of the conditional variance. If an ARCH process is the true data generating mechanism, the nonlinearities cannot be exploited to generate improved point predictions relative to linear model.

In the sequel, the conditional mean is considered as an l^{th} order autoregressive process defined by

$$E(y_{t+l}) = c_0 + \sum_{i=1}^l c_i y_{t-i} \quad (2)$$

assuming the unpredictable component in (1) is an ARCH process, it can be represented as:

$$\begin{aligned} \varepsilon_t &= z_t \sigma_t \\ z_t &\stackrel{iid}{\sim} N(0,1) \\ \sigma_t^2 &= f(\sigma_{t-1}^2, \sigma_{t-2}^2, \dots, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, x_{t-1}), \end{aligned} \quad (3)$$

here z_t is an independently and identically distributed i.d. process, $E(z_t) = 0$, $V(z_t) = 1$ and σ_t is a time-varying, positive measurable function of the information set time $t-1$. The unpredictable component has variance σ_t^2 , conditional on information given at time $t-1$. The conditional variance is a linear or nonlinear function of lagged conditional variances, past prediction errors,

exogenous and endogenous variables measurable at time $t-1$. The conditional prediction error is normally distributed, but the unconditional prediction error and the conditional variance of it have an unknown form of distribution. The conditional standardized prediction error, z_{t+l} , is standard normally distributed:

$$\varepsilon_{t+l} \sim N(0, \sigma_t^2) \Leftrightarrow z_{t+l} \equiv \varepsilon_{t+l} \sigma_t^{-1} \sim N(0,1). \quad (4)$$

In the recent literature, one can find a vast number of parametric specifications of ARCH models motivated by the characteristics explored in financial markets. A researcher, who is looking for the "best" model, would have in mind a variety of candidate models. The most commonly used conditional variance functions are the GARCH (Bollerslev [4]), the Exponential GARCH, or E-GARCH, (Nelson [29]) and the Threshold GARCH, or TARCH, (Glosten et al. [16]) functions. In the sequel, these ARCH models are considered in the following forms:

The GARCH(p,q) model

$$\sigma_t^2 = a_0 + \sum_{i=1}^q (a_i \varepsilon_{t-i}^2) + \sum_{i=1}^p (b_i \sigma_{t-i}^2) \quad (5)$$

The E-GARCH(p,q) model

$$\ln(\sigma_t^2) = a_0 + \sum_{i=1}^q \left(\alpha_i \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| + \gamma_i \left(\frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right) \right) + \sum_{i=1}^p (b_i \ln(\sigma_{t-i}^2)) \quad (6)$$

The TARCH(p,q) model

$$\sigma_t^2 = a_0 + \sum_{i=1}^q (c_i \varepsilon_{t-i}^2) + \gamma \varepsilon_{t-1}^2 d_{t-1} + \sum_{i=1}^p (b_i \sigma_{t-i}^2), \quad (7)$$

where $d_t = 1$ if $\varepsilon_t < 0$, and $d_t = 0$ otherwise.

Maximum likelihood estimates of the parameters are obtained by numerical maximization of the log-likelihood function using the Marquardt algorithm (Marquardt [28]), a modification of the Berndt, Hall, Hall and Hausman, or BHHH, algorithm (Berndt et al. [3]). The quasi-maximum likelihood estimator (QMLE) is used, as according to Bollerslev and Wooldridge [5], it is generally consistent, has a normal limiting distribution and provides asymptotic standard errors that are valid under non-normality.

The majority of practical applications, i.e. option pricing, determination of the value-at-risk, require more than one-day-ahead volatility forecasts. More than one-step-ahead forecasts can be computed by repeated substitution. The forecast recursion relation of the GARCH(p,q) model is:

$$\begin{aligned} \sigma_{t+l|t}^2 &= a_0^{(l)} + \sum_{i=1}^q (a_i^{(l)} \varepsilon_{t-i+1|t}^2) + \sum_{i=1}^p (b_i^{(l)} \sigma_{t-i+1|t}^2) \\ \sigma_{t+s|t}^2 &= a_0^{(s)} + \sum_{i=1}^q (a_i^{(s)} \sigma_{t-i+1|t}^2) + \sum_{i=1}^p (b_i^{(s)} \sigma_{t-i+1|t}^2) \end{aligned} \quad (8)$$

²for an overview of the Neural Networks literature see Poggio and Girosi [32], Hertz et al. [19], White [41], Hutchinson et al. [24].

³Brock [8], Holden [22], Thompson and Stewart [38] and Sieh [23] review applications of chaotic systems to financial markets.

⁴Priestley [33], Tong [39] and Teräsvirta et al. [37] cover a wide variety of nonlinear models.

⁵for an overview of the ARCH literature see Bollerslev et al. [6], [7], Bera and Higgins [2], Hamilton [18], Gouriéroux [17].

see for details Taylor [36] and Shephard [35].

For $s > t$, the forecast of predictive error ε_s conditional on information available at time t equals to its zero expected value, $E(\varepsilon_s | I_t) = 0$. On the other hand, the estimated value of ε_t^2 measured at time t should be equal to σ_{tt}^2 for $s > t$. For $s \leq t$, the predictive error and its square are computed by the model with the available information at time t . The forecast recursion relationship associated with the E-GARCH(p,q) model is:

$$\begin{aligned} \ln(\sigma_{t+1|t}^2) &= a_0^{(t)} + \sum_{i=1}^q \left(a_i^{(t)} \left| \frac{\varepsilon_{t-i+1|t}}{\sigma_{t-i+1|t}} \right| + \gamma_i^{(t)} \left(\frac{\varepsilon_{t-i+1|t}}{\sigma_{t-i+1|t}} \right)^2 \right) + \\ &\quad + \sum_{i=1}^p (b_i^{(t)} \ln(\sigma_{t-i+1|t}^2)) \\ \ln(\sigma_{t+s|t}^2) &= a_0^{(t)} + \sum_{i=s}^q \left(a_i^{(t)} \left| \frac{\varepsilon_{t-i+1|t}}{\sigma_{t-i+1|t}} \right| + \gamma_i^{(t)} \left(\frac{\varepsilon_{t-i+1|t}}{\sigma_{t-i+1|t}} \right)^2 \right) + \\ &\quad + \sqrt{\frac{2}{\pi}} \sum_{i=1}^q (a_i^{(t)}) + \sum_{i=1}^p (b_i^{(t)} \ln(\sigma_{t-i+1|t}^2)) \end{aligned} \quad (9)$$

that associated with the TAR(p,q) model is:

$$\begin{aligned} \sigma_{t+1|t}^2 &= a_0^{(t)} + \sum_{i=1}^q (c_i^{(t)} \varepsilon_{t-i+1|t}^2) + \gamma^{(t)} \varepsilon_{tt}^2 d_t + \\ &\quad + \sum_{i=1}^p (b_i^{(t)} \sigma_{t-i+1|t}^2) \\ \sigma_{t+s|t}^2 &= a_0^{(t)} + \sum_{i=s}^q (c_i^{(t)} \varepsilon_{t-i+1|t}^2) + \sum_{i=1}^q (c_i^{(t)} \varepsilon_{t-i+1|t}^2) + \\ &\quad + \gamma^{(t)} \sigma_{t-i+1|t}^2 E(d_t) + \sum_{i=1}^p (b_i^{(t)} \sigma_{t-i+1|t}^2) \end{aligned} \quad (10)$$

$E(d_t)$ denotes the percentage of negative innovations out of all innovations. Under the assumption of normally distributed innovations, the expected number of negative shocks is equal to the expected number of positive shocks, or $E(d_t) = 0.5$.

The forecast of the conditional variance at time t over a horizon of N days ahead is simply the average of the estimated future variance conditional on information given at time t :

$$\sigma_{t+N}^2 = N^{-1} \sum_{i=1}^N \sigma_{t+i|t}^2 \quad (11)$$

III. The Correlated Gamma Ratio (CGR) Model Selection Method

Degianmakis and Xekalaki [12] compare the forecasting ability of ARCH models using the Correlated Gamma Ratio (CGR) distribution. This is a distribution derived by Panaretos et al. [31] as the distribution of the ratio of two variables jointly distributed according to

Kibble's [25] Bivariate Gamma distribution. Kibble [25] proves that if, for $t=1,2,\dots$, the joint distribution of $(r_t^{(A)}, r_t^{(B)})$ is the Bivariate Standard Normal, then the joint

distribution of $T^{-1} \sum_{t=1}^T r_t^{(A)2}$ and $T^{-1} \sum_{t=1}^T r_t^{(B)2}$ is Kibble's

Bivariate Gamma distribution. As pointed out by Panaretos et al. [31], $r_t^{(A)}$ and $r_t^{(B)}$ could represent the standardized prediction errors from two regression models (not necessarily nested) but with a common dependent variable. The distribution of the ratio of the sum of their squares is the Correlated Gamma Ratio distribution,

$\sum_{t=1}^T r_t^{(B)2} / \sum_{t=1}^T r_t^{(A)2} \sim CGR(k, \rho)$, where $k = T/2$ and $\rho = \text{Cor}(r_t^{(A)}, r_t^{(B)})$. Thus, two regression models can be

compared through testing the null hypothesis of equivalence of the models in their predictability against the alternative that model (A) produces "better" predictions. The null

hypothesis is rejected if $\sum_{t=1}^T r_t^{(B)2} / \sum_{t=1}^T r_t^{(A)2} > CGR(k, \rho, \alpha)$,

where $CGR(k, \rho, \alpha)$ is the $1-\alpha$ percentile of the CGR distribution⁷.

Let us now assume that we are interested in comparing the predictive ability of two ARCH models:

Model A
$\varepsilon_t^{(A)} = z_{1,t} \sigma_t^{(A)}$
$z_{1,t} \sim N(0,1)$
$\sigma_t^{(A)2} = f(\sigma_{t-1}^{(A)2}, \dots, \sigma_{t-p}^{(A)2}, \varepsilon_{t-1}^{(A)2}, \dots, \varepsilon_{t-q}^{(A)2}, x_{t-1}^{(A)})$
Model B
$\varepsilon_t^{(B)} = z_{2,t} \sigma_t^{(B)}$
$z_{2,t} \sim N(0,1)$
$\sigma_t^{(B)2} = f(\sigma_{t-1}^{(B)2}, \dots, \sigma_{t-p}^{(B)2}, \varepsilon_{t-1}^{(B)2}, \dots, \varepsilon_{t-q}^{(B)2}, x_{t-1}^{(B)})$

The joint distribution of $T^{-1} \sum_{t=1}^T z_{t|t-1}^{(A)2} \equiv T^{-1} \sum_{t=1}^T \varepsilon_{t|t-1}^{(A)2} / \sigma_t^{(A)2}$

and $T^{-1} \sum_{t=1}^T z_{t|t-1}^{(B)2} \equiv T^{-1} \sum_{t=1}^T \varepsilon_{t|t-1}^{(B)2} / \sigma_t^{(B)2}$ is Kibble's Bivariate

Gamma distribution. Thus, the standardized one-step-ahead prediction errors can be used to test the null hypothesis of equivalence of the models in their predictive ability against the alternative that the first model produces "better" predictions. The null hypothesis is rejected if

$$\sum_{t=1}^T z_{t|t-1}^{(B)2} / \sum_{t=1}^T z_{t|t-1}^{(A)2} > CGR(k, \rho, \alpha).$$

According to the CGR model selection method, the models that are considered as having a better ability to predict future values of the dependent variable, are those with the lowest sum of squared standardized one-step-ahead

⁷ Percentage points of the CGR distribution can be found in Panaretos et al. [31].

prediction errors. It becomes evident, therefore, that these models can potentially be regarded as the most appropriate to use for volatility forecasts too.

Let us assume that M candidate ARCH models are available and that we are looking for the "most suitable" model at each of a sequence of points in time. At time k , selecting a strategy for the most appropriate model to forecast volatility at time $k+1$ ($k=T, T+1, \dots$) could naturally amount to selecting the model which, at time k , has the lowest sum of squared standardized one-step-ahead prediction errors, on the basis of the CGR criterion. Table I summarizes the estimation steps comprising this approach.

Model	Time		
	$k=T$	$k=T+1$	$k=T+K$
$m=1$	$\sum_{t=1}^T z_{t T-1}^{(1)2}$	$\sum_{t=2}^{T+1} z_{t T-1}^{(1)2}$	$\sum_{t=T+K+1}^{T+K} z_{t T-1}^{(1)2}$
$m=2$	$\sum_{t=1}^T z_{t T-1}^{(2)2}$	$\sum_{t=2}^{T+1} z_{t T-1}^{(2)2}$	$\sum_{t=T+K+1}^{T+K} z_{t T-1}^{(2)2}$
$m=M$	$\sum_{t=1}^T z_{t T-1}^{(M)2}$	$\sum_{t=2}^{T+1} z_{t T-1}^{(M)2}$	$\sum_{t=T+K+1}^{T+K} z_{t T-1}^{(M)2}$

Table I
The estimation steps required at time k for each model m by the CGR model selection method. (At time k ($k=T, T+1, \dots$), select the model m with minimum

$$\sum_{t=k-T+1}^k z_{t|T-1}^{(m)2})$$

In the next section, the methodology applied to evaluate the performance of a model in estimating future volatility is presented, while in section V, the ability of the CGR model selection method to indicate those ARCH models that generate "better" volatility predictions is illustrated on a set of real data on daily returns of the S&P500 stock index.

IV. Evaluating the Volatility Forecast Performance

The main problem in evaluating the predictive performance of a model is the choice of the function one should use to measure the distance between estimations and observations. Evaluating the performance of the variance forecasts requires knowledge of the actual volatility, which is unobservable. Thus, in evaluating the predictive performance of a variance model a question of a dual nature arises: that of determining the realized volatility and of considering the appropriate measure to evaluate the closeness of the forecasts to the corresponding realizations.

A Realized Volatility Measures

Practitioners' most popular volatility measures are the average of squared daily returns and the variance of the

daily returns. These measures, expressed on a daily basis for a horizon of N days ahead, are:

$$\bar{s}_{d(N)}^2 = N^{-1} \sum_{t=1}^N v_{t-t}^2, \quad (12)$$

$$\bar{s}_{d(N)}^2 = (N-1)^{-1} \sum_{t=1}^N (v_{t+1} - \bar{y}_{d(N)})^2, \quad (13)$$

respectively, where $\bar{y}_{d(N)} = N^{-1} \sum_{t=1}^N v_{t+1}$ is the average return.

B Evaluation Criteria

A large number of forecast evaluation criteria exist in the literature. However, none is generally acceptable. Because of high non-linearity in volatility models and the variety of statistical evaluation criteria, a number of researchers constructed economic criteria based upon the goals of their particular application. West et al. [40] develop a criterion based on the decisions of a risk averse investor. Engle et al. [15] assume that the objective is to price options and develop a loss function from the profitability of a particular trading strategy. In the sequel, we focus on statistical criteria to measure the closeness of the forecasts to the realizations, in order to avoid restrictions imposed by economic theory. Moreover, we consider statistical criteria that are robust to non-linearity and heteroscedasticity. Pagan and Schwert [30] use statistical criteria to compare parametric and non-parametric ARCH models with in-sample and out-of-sample data. Besides, Heynen and Kat [20] investigate the predictive performance of ARCH and Stochastic Volatility models and Hol and Koopman [21] compare the predictive ability of Stochastic Volatility and Implied Volatility models. Andersen et al. [1] applied heteroscedasticity-adjusted statistics to examine the forecasting performance of intraday returns. Denoting the forecasting variance over an N day period measured at day t by $\sigma_{d(N)}^2$, and the realized variance over the same period by $s_{d(N)}^2$, the following evaluation criteria are considered:

Squared Error (SE):

$$(\sigma_{d(N)}^2 - s_{d(N)}^2)^2 \quad (14)$$

Absolute Error (AE):

$$|\sigma_{d(N)}^2 - s_{d(N)}^2| \quad (15)$$

Heteroscedasticity Adjusted Squared Error (HASE):

$$(1 - s_{d(N)}^2 / \sigma_{d(N)}^2)^2 \quad (16)$$

Heteroscedasticity Adjusted Absolute Error (HAAE):

$$|1 - s_{d(N)}^2 / \sigma_{d(N)}^2| \quad (17)$$

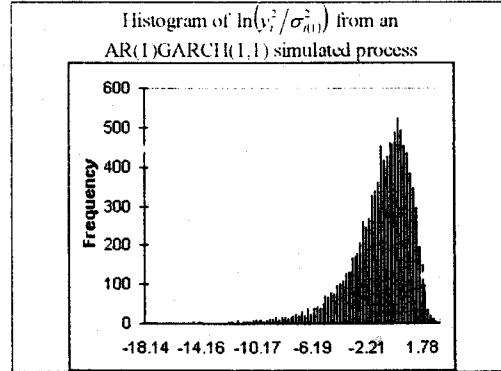
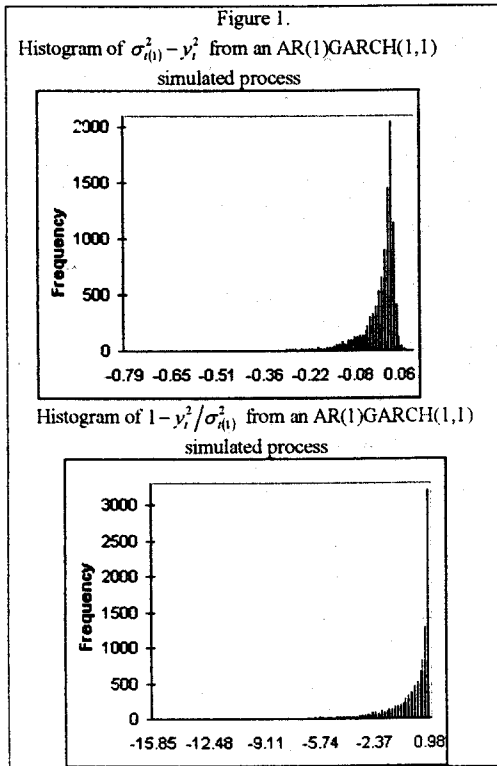
Logarithmic Error (LE):

$$\ln(s_{d(N)}^2 / \sigma_{d(N)}^2)^2 \quad (18)$$

The first two functions have been widely used in the literature (see, e.g. Heynen and Kat [20]). The HASE and HAAE functions were considered by Andersen et al. [1], while the LE function was utilized by Pagan and Schwert [30].

Usually, the average of the evaluation criteria is considered. However, when simulating an AR(1)GARCH(1,1) process, which is the most commonly used model in financial applications, the distributions of $(\sigma_{t(N)}^2 - s_{t(N)}^2)$, $(1 - s_{t(N)}^2/\sigma_{t(N)}^2)$ and $\ln(s_{t(N)}^2/\sigma_{t(N)}^2)$ are asymmetric with extreme outliers. It would therefore be advisable to compute both the mean and the median of the evaluation criteria. Figure 1 depicts the histograms of the one-step forecast error distribution from the following simulated process:

$$\begin{aligned} y_t &= 0.001 + 0.1y_{t-1} + \varepsilon_t \\ \sigma_t^2 &= 0.002 + 0.05\varepsilon_{t-1}^2 + 0.9\sigma_{t-1}^2 \\ \varepsilon_t &= z_t\sigma_t, \text{ and } z_t \sim N(0,1) \end{aligned} \quad (19)$$



V. Examining the Performance of the CGR Model Selection Method

In this section, the ability of the CGR model selection method to lead to the ARCH models that track closer future volatility is illustrated on a real stock index daily return series. As follows from section II, the return series can be modeled in the following form:

$$y_t = E(y_{t+1}) + \varepsilon_t$$

$$E(y_{t+1}) = c_0 + \sum_{i=1}^4 c_i y_{t-i} \quad (20)$$

$$\varepsilon_t = z_t \sigma_t$$

$$z_t \sim N(0,1)$$

$$\sigma_t^2 = f(\sigma_{t-1}^2, \dots, \sigma_{t-p}^2, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2, x_{t-1})$$

In the sequel, the above form is considered in connection with the ARCH models defined by (5), (6) and (7), for $l=0,1,2,3,4$, $p=0,1,2$ and $q=1,2$, thus yielding a total of 85 cases⁸.

The data set consists of 1661 S&P500 stock index daily returns in the period from November 24th, 1993 to June 6th, 2000. The ARCH processes are estimated using roll sample of constant size equal to 500. Thus, the first one-step-ahead volatility prediction, $\sigma_{t+1|t}^2$, is available at time $t=500$. Applying the CGR model selection method, the sum of squared standardized one-step-ahead prediction errors, $\sum_{t=500}^T z_{t+1}^2$, was estimated considering various values for T , and, in particular, $T=5(5)80$. This is an indirect way to examine the performance of the CGR model selection method for various values of T . Thus, the evaluation criteria were applied on the one-step-ahead forecasts using $1661-500-80=1081$ data points, on the two-step-ahead forecasts using $1661-500-81=1080$ data

⁸ Numerical maximization of the log-likelihood function, for the E-GARCH(2,2) model, frequently failed to converge. So the five E-GARCH models for $p=q=2$ were excluded.

points and on the k^{th} -step-ahead forecasts using $1081-k+1$ data points.

Our main purpose is to examine the application potential of the CGR method of selection of models on the basis of their forecasting ability in terms of volatility. So, the mean and the median value of each of the 5 evaluation criteria, in equations (14)-(18), were computed, yielding a total of 10 evaluation criteria for each forecasting horizon from one day to one hundred days ahead. However, volatility is expressed either as the variance or as the standard deviation of the returns. Thus, in order to examine possible differences between forecasting the variance and its square root, the evaluation criteria were, also, applied on the standard deviation. Therefore, $\sigma_{i(N)}^2$ and $s_{i(N)}^2$, in equations (14)-(18), were replaced by $\sigma_{i(N)}$ and $s_{i(N)}$, respectively and 10 more evaluation criteria were computed. In total, 20 evaluation criteria were computed for a horizon ranging from one day to five months. In section IV.A, two realized volatility measures were mentioned. The results are based on the realized volatility as defined by $\hat{s}_{i(N)}^2$. (Results based on the realized volatility as defined by $\hat{s}_{i(N)}^2$ are similar and are not reported.)

It was examined whether the ARCH models selected by the CGR method achieve the lowest value of the evaluation criteria. The main focus was on the median values of the criteria and mainly on the heteroscedasticity adjusted criteria since they are more robust to asymmetry. Figure 2 shows, for each evaluation criterion and each forecasting horizon, whether ARCH models selected by the CGR method achieve the lowest value of the evaluation criteria. In the first part of Figure 2, the performance of the models, which are selected by the CGR method, on the basis of the conditional variance is depicted, while, the second part refers to their performance on forecasting standard deviation. The general conclusion is that the CGR method led to the selection of the ARCH processes which track closer the realized volatility in the majority of the cases. Specifically, for the forecasting horizon ranging from 11 to 52 days, the models selected by the CGR method achieve the lowest criteria values, irrespectively the evaluation criteria. The percentage of cases, that the models selected by the CGR method achieve the lowest value of the evaluation criteria, is higher around the forecasting horizon ranging from 16 to 36 days ahead, or 4 to 7 trading weeks ahead. The ability of the CGR method to select the ARCH models that generate "better" predictions of the volatility, around a forecasting horizon of 4 to 7 weeks ahead, is indicative of its usage potential in applications exploiting volatility forecasts as, for example, in pricing derivatives, estimating the risk of a portfolio etc. Table II presents the percentage of cases the models selected by the CGR method perform "better" as judged by the evaluation criteria, for 3 different horizon ranges. Note that, in terms of the MSE and MAE criteria, none of the models chosen by the CGR method appears to perform better in any of the forecasting horizons considered. But, in terms of the median values of the criteria and the heteroscedasticity adjusted criteria,

which are robust to asymmetry, the models selected by the CGR method appear to have a better performance in all the forecasting horizons considered.

It is interesting to note that, via the evaluation criteria considered, the optimum sample size, T , for the CGR model selection method can be determined. The CGR model selection method has been applied for $T = 5(5)80$. In the sequel, the value of T for which the CGR selection method achieves the best performance according to the evaluation criteria used, is examined. Figure 3 shows a plot of the average T , suggested by the evaluation criteria, across the forecasting horizons. The bar charts of Figure 3 are a graphical representation of the number of evaluation criteria by which the performance of the models selected by the CGR method were judged "best" (measured on the right hand side vertical axis).

For a 16 to 36 day ahead forecasting horizon, the appropriate T , as concerns the specific data, ranges around 20 days with a standard deviation of 3.6 days. Table III provides more details for the sample size of the CGR selection method suggested by the evaluation criteria and its standard deviation for both the entire 16 to 36 day ahead forecasting horizon and for each day individually. The CGR model selection method shows a better performance for a sample size of about 20 days.

In order to test the importance of selecting the appropriate T , for the model selection method suggested, the evaluation criteria were run for $T = 5(5)80$. The results are indeed in support of a sample size of around 20 days for the CGR method to manifest a better performance. Figure 4 presents the percentage of the evaluation criteria by which the CGR method, with specific T , selects those ARCH models that generate "better" volatility predictions. For T ranging from 15 to 35, the CGR selection method appears to have the highest performance.

VI. Conclusions

A method for selecting an ARCH model among several competing models was suggested based on the CGR distribution. It amounts to selecting the model with the lowest sum of squared standardized forecasting errors. A number of evaluation criteria, for forecasting horizons ranging from one day to one hundred days ahead, were applied and it was found that the ARCH models, selected by the CGR model selection method, generate "better" predictions of the volatility. Thus, the CGR selection method appears to be a useful tool in guiding one's choice of the appropriate model for estimating future volatility, with applications in evaluating portfolios, derivatives and financial risk.

References

- [1] Andersen, T., T. Bollerslev and S. Lange (1999). Forecasting Financial Market Volatility: Sample Frequency vis-à-vis Forecast Horizon. *Journal of Empirical Finance*, 6, 457-477.
- [2] Bera, A.K., and M.L. Higgins (1993). ARCH Models: Properties, Estimation and Testing. *Journal of Economic Surveys*, 7, 305-366.

- [3] Berndt, E., B. Hall, R. Hall, and J. Hausman. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, 3, 653-665.
- [4] Bollerslev, T., (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-327.
- [5] Bollerslev, T., and J. Wooldridge. (1992). Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances. *Econometric Reviews*, 11, 143-172.
- [6] Bollerslev, T., R. Chou, and K. Kroner. (1992). ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, 52, 5-59.
- [7] Bollerslev, T., R. Engle, and D. Nelson. (1994). ARCH Models. in *Handbook of Econometrics*, Volume 4, eds. R. Engle and D. McFadden. Elsevier Science, Amsterdam, 2959-3038.
- [8] Brock, W., (1986). Distinguishing Random and Deterministic Systems: Abridged Version. *Journal of Economic Theory*, 40, 168-195.
- [9] Brock, W., W. Dechert, and J. Scheinkman. (1987). A Test for Independence Based on the Correlation Dimension. *Department of Economics, University of Wisconsin, Madison, WI. SSRI, Working Paper no. 8702*.
- [10] Campbell, J., A. Lo, and A. C. MacKinlay. (1997). *The Econometrics of Financial Markets*, New Jersey: Princeton University Press.
- [11] Cohen, K., G. Hawawini, S. Maier, R. Schwartz, and D. Whitcomb. (1983). Friction in the Trading Process and the Estimation of Systematic Risk. *Journal of Financial Economics*, 12, 263-278.
- [12] Degiannakis, S., and E. Xekalaki. (1999). Predictability and Model Selection in the Context of ARCH Models. *Athens University of Economics and Business, Department of Statistics*, Technical Report, 69.
- [13] Degiannakis, S., and E. Xekalaki. (2001). Measuring the Predictive Accuracy of the Correlated Gamma Ratio Method for Model Selection. *Athens University of Economics and Business, Department of Statistics*, Technical Report, 134.
- [14] Dimson, E., (1979). Risk Measurement when Shares are Subject to Infrequent Trading. *Journal of Financial Economics*, 7, 197-226.
- [15] Engle, R., C. Hong, A. Kane, and J. Noh. (1993). Arbitrage Valuation of Variance Forecasts with Simulated Options. *Advances in Futures and Options Research*, 6, 393-415.
- [16] Glosten, L., R. Jagannathan, and D. Runkle. (1993). On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48, 1779-1801.
- [17] Gouriéroux C., (1997). *ARCH models and Financial Applications*, Springer Series.
- [18] Hamilton, J., (1994). *Time Series Analysis*, New Jersey: Princeton University Press.
- [19] Hertz, J., A. Krogh, and R. Palmer. (1991). Introduction to the Theory of Neural Computation. *Addison-Wesley Publishing Company*, Reading, MA.
- [20] Heynen, R., and H. Kat. (1994). Volatility Prediction: A Comparison of the Stochastic Volatility, GARCH(1,1), and EGARCH(1,1) Models. *Journal of Derivatives*, Winter 94, 50-65.
- [21] Hol, E., and S. Koopman. (2000). Forecasting the Variability of Stock Index Returns with Stochastic Volatility Models and Implied Volatility. *Timbergen Institute*, Discussion Paper No. 104.4.
- [22] Holden, A., (1986). *Chaos*, New Jersey: Princeton University Press.
- [23] Hsieh, D., (1991). Chaos and Nonlinear Dynamics: Application to Financial Markets. *Journal of Finance*, 46, 1839-1877.
- [24] Hutchinson, J., A. Lo, and T. Poggio. (1994). A Nonparametric Approach to the Pricing and Hedging of Derivative Securities Via Learning Networks. *Journal of Finance*, 49, 851-889.
- [25] Kibble, W. F., (1941). A Two-Variate Gamma Type Distribution. *Sankhya*, 5, 137-150.
- [26] Lo, A., and A. C. MacKinlay. (1988). Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. *Review of Financial Studies*, 1, 41-66.
- [27] Lo, A., and A. C. MacKinlay. (1990). An Econometric Analysis of Non-synchronous Trading. *Journal of Econometrics*, 45, 181-212.
- [28] Marquardt, D. W., (1963). An Algorithm for Least Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11, 431-441.
- [29] Nelson, D., (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59, 347-370.
- [30] Pagan, A., and G. Schwert. (1990). Alternative Models for Conditional Stock Volatility. *Journal of Econometrics*, 45, 267-290.
- [31] Panaretos, J., S. Psarakis, and E. Xekalaki. (1997). The Correlated Gamma Ratio Distribution in Model Evaluation and Selection. *Athens University of Economics and Business, Department of Statistics*, Technical Report, 33.
- [32] Poggio, T., and F. Girosi. (1990). Networks for Approximation and Learning. *Proceeding of the IEEE*, special issue: Neural Networks I: Theory and Modeling, 78, 1481-1497.
- [33] Priestley, M., (1988). *Nonlinear and Non-Stationary Time Series Analysis*, Academic Press, San Diego.
- [34] Scholes, M., and J. Williams. (1977). Estimating Betas from Nonsynchronous Data. *Journal of Financial Economics*, 5, 309-328.
- [35] Shephard, N. (1995). Statistical Aspects of ARCH and Stochastic Volatility. *University of Oxford, Nuffield College, Discussion Paper 94*.
- [36] Taylor, S.J. (1994). Modeling Stochastic Volatility. *Mathematical Finance*, 4, 2, 183-204.
- [37] Teräsvirta, T., D. Tjøstheim, and C. Granger. (1994). Aspects of Modeling Nonlinear Time Series, in *Handbook of Econometrics*, Volume 4, eds. R. Engle and D. McFadden, Elsevier Science, Amsterdam.
- [38] Thompson, J., and H. Stewart. (1986). *Nonlinear Dynamics and Chaos*, John Wiley and Sons, NY.
- [39] Tong, H., (1990). *Nonlinear Time Series: A Dynamic System Approach*, Oxford University Press, Oxford.
- [40] West, K. D., H. J. Edison, and D. Cho. (1993). A Utility Based Comparison of Some Models for Exchange Rate Volatility. *Journal of International Economics*, 35, 23-45.
- [41] White, H., (1992). *Artificial Neural Networks: Approximation and Learning Theory*, Blackwell Publishers, Cambridge, MA.

