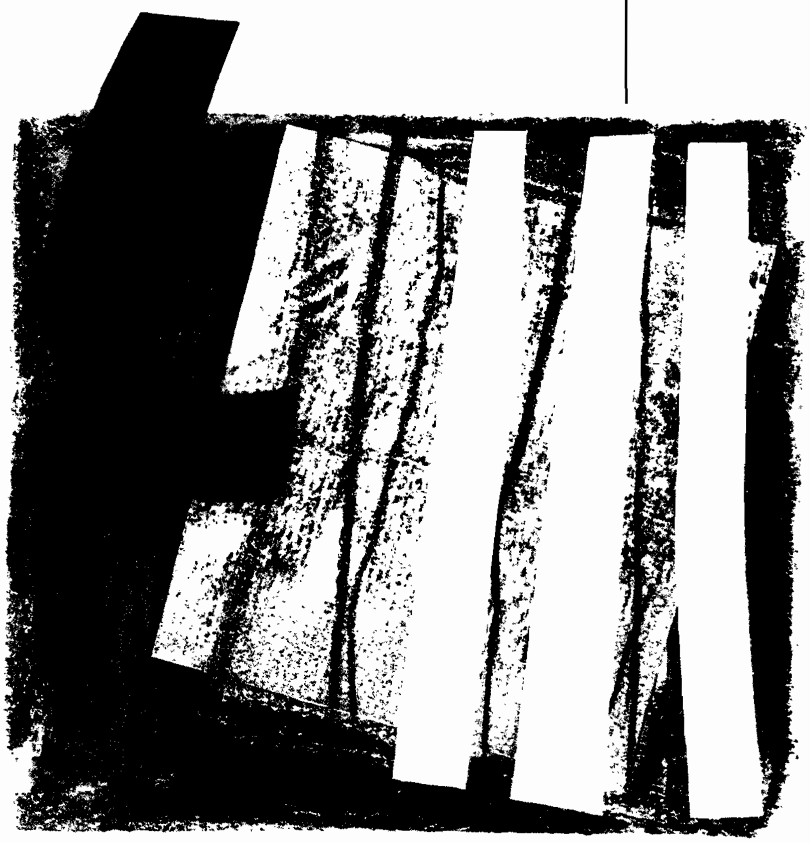
FORECASTING RETURNS AND VOLATILITIES IN GARCH PROCESSES USING THE BOOTSTRAP

Lorenzo Pascual, Juan Romo and Esther Ruiz

00-68



WORKING PAPERS

Universidad Carlos III de Madrid

Working Paper 00-68
Statistics and Econometrics Series 31
2000

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 624-9849

FORECASTING RETURNS AND VOLATILITIES IN GARCH PROCESSES USING THE BOOTSTRAP

Lorenzo Pascual, Juan Romo and Esther Ruiz*

Abstract

We propose a new bootstrap resampling scheme to obtain prediction densities of levels and volatilities of time series generated by GARCH processes. The main advantage over other bootstrap methods previously proposed for GARCH processes, is that the procedure incorporates the variability due to parameter estimation and, consequently, it is possible to obtain bootstrap prediction densities for the volatility process. The asymptotic properties of the procedure are derived and the finite sample properties are analysed by means of Monte Carlo experiments, showing its good behaviour versus alternative procedures. Finally, the procedure is applied to estimate prediction densities of returns and volatilities of the Madrid Stock Market index, IBEX-35.

Keywords: Time series; Forecasting; Non gaussian distributions; Non linear models; Resampling methods.

*Pascual, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Madrid. Spain. Ph: 34-91-624.98.59, Fax: 34-91-624.98.49, e-mail: pascual@est-econ.uc3m.es; Romo, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, Ph: 34-91-624.98.05, e-mail: romo@est-econ.uc3m.es; Ruiz, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, Ph: 34-91-624.98.56, e-mail: ortega@est-econ.uc3m.es

Forecasting returns and volatilities in GARCH processes using the bootstrap

Lorenzo Pascual, Juan Romo and Esther Ruiz
Departamento de Estadística y Econometría
Universidad Carlos III de Madrid

Abstract

We propose a new bootstrap resampling scheme to obtain prediction densities of levels and volatilities of time series generated by GARCH processes. The main advantage over other bootstrap methods previously proposed for GARCH processes, is that the procedure incorporates the variability due to parameter estimation and, consequently, it is possible to obtain bootstrap prediction densities for the volatility process. The asymptotic properties of the procedure are derived and the finite sample properties are analysed by means of Monte Carlo experiments, showing its good behaviour versus alternative procedures. Finally, the procedure is applied to estimate prediction densities of returns and volatilities of the Madrid Stock Market index, IBEX-35.

KEY WORDS: Time series, Forecasting, Non Gaussian distributions, Non linear models, Resampling methods.

1. INTRODUCTION

It is by now well documented in the literature that high frequency financial time series are characterized by having conditional heteroscedasticity. Generalized Autoregressive Conditionally Heteroscedastic (GARCH) models were originally introduced by Engle (1982) and Bollerslev (1986) to represent the dynamic evolution of conditional variances. One of the motivations of GARCH models was to provide dynamic prediction intervals with the intervals being narrow in tranquil times and wide in volatile periods. Furthermore, financial market participants have shown an increasing interest in interval forecast as measures of uncertainty. For example, in the area of financial risk management, it is of interest to provide density forecasts of portfolio prices and to track certain aspects of these densities such as value at risk. However, despite the extensive literature related with GARCH models, relatively little attention has been given to the construction of prediction intervals of GARCH models. One step ahead prediction errors of conditionally Gaussian GARCH models are Normally distributed but the distribution of prediction errors more than one step ahead is unknown. Baillie and Bollerslev (1992) use Cornish-Fisher expansions to obtain prediction intervals of nonlinear regression functions with ARMA disturbances and GARCH(1,1) innovations, making parametric hypothesis on the conditional distribution of the innovations. However, generalizations for other non Gaussian GARCH models are not available.

On the other hand, the volatility of returns is a key factor in many models of option valuation and portfolio allocation problems. Therefore, accurate predictions of volatilities are critical for the implementation and evaluation of asset and derivative pricing theories as well as trading and hedging strategies. However, the literature on volatility prediction has only deal with point forecasts without giving any measure of prediction uncertainty when forecasting future volatilities. For example, Andersen and Bollerslev (1998) and Andersen et al. (1999) explore the return volatility predictability inherent in high-frequency returns finding that both model misspecification and parameter estimation error should detract from the predictive power of GARCH models. However, there are not results on the distribution of prediction intervals for future volatilities.

Bootstrap based methods allow to obtain prediction intervals that incorporate the uncertainty due to parameter estimation without distributional assumptions on the sequence of innovations. In the context of linear time series, different bootstrap methods have been proposed in the literature to improve the prediction intervals based on Box and Jenkins (1976). The procedure originally proposed by Thombs and Schucany (1990) try to estimate directly the distribution of the prediction k-periods ahead of AR(p) models, conditional on the information given by the observed data. This approach needs the backward representation of the autoregressive model to construct bootstrap replicates that mimic the structure of the original data. The need of the backward representation makes this method computationally expensive and, what is more important, restrict its applicability to models having this representation, excluding GARCH processes. Cao *et al.* (1997) present an alternative bootstrap method for constructing prediction intervals for stationary AR(p) models which does not require the backward representation. However, their intervals do not incorporate the variability due to parameter estimation. Miguel and Olave (1999) extend the procedure of Cao *et al.* (1997) to construct prediction intervals for ARMA processes with ARCH innovations. However, their proposal do not allow to construct prediction intervals for future volatilities.

More recently, Pascual, Romo and Ruiz (1998) have proposed a bootstrap procedure for autoregressive integrated moving average (ARIMA) processes that is able to take into account the uncertainty due to parameter estimation and that does not require resampling through the backward representation of the process. In this paper, we generalize the procedure of Pascual, Romo and Ruiz (1998) to estimate prediction densities of both, returns and volatilities generated by GARCH models.

The paper is organized as follows. Section 2 describes the main properties of GARCH processes and predictions. In section 3, we present the proposed resampling procedure to estimate prediction densities and intervals for returns and volatilities. In section 4, we derive the asymptotic properties of the proposed bootstrap procedure. Its finite sample behavior is analyzed in section 5 that reports results of an extensive Monte Carlo simulation study. Section 6 presents an application with real financial data. Finally, the conclusions and some ideas for future research appear in Section 7.

2. THE GARCH(1,1) MODEL

The GARCH(1,1) model provides a simple representation of the main statistical characteristics of return series of a wide range of assets and, consequently, it is extensively used to model real financial time series. Hence, although it could not be the optimal model for volatility forecasting in any given series, it does serve as a natural benchmark for the forecast performance of heteroscedastic models based on ARCH. In the simplest set up, a GARCH(1,1) model is given by

$$\begin{aligned} y_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, t = 1, ..., T \end{aligned} \tag{1}$$

where ε_t is a white noise process with unity variance, σ_t is a stochastic process known as volatility that it is assumed to be independent of ε_t and ω, α and β are unknown parameters that satisfy $\omega > 0$, $\alpha, \beta \geq 0$ to ensure the positivity of the conditional variance. The process y_t is stationary if $\alpha + \beta < 1$. Notice that σ_t^2 is observable with information available at time t-1 and, consequently, given the assumptions on the distribution of ε_t, the conditional mean of y_t is zero and σ_t^2 is the conditional variance. Further, the conditional distribution of y_t coincides with the distribution of ε_t.

Alternatively, the conditional variance can also be expressed as a function of past observations as follows:

$$\sigma_t^2 = \frac{\omega}{1-\alpha-\beta} + \alpha \sum_{j=0}^{\infty} \beta^j (y_{t-j-1}^2 - \frac{\omega}{1-\alpha-\beta}). \tag{2}$$

Although the marginal distribution of y_t is, in general, unknown, it is easy to prove that the distribution of y_t has thick tails with zero mean and if $\alpha+\beta < 1$, the marginal variance is given by:

$$E(y_t^2) = \frac{\omega}{1-\alpha-\beta}. \tag{3}$$

Finally, the predictor of y_{T+k} given observations of the process up to time T, $\{y_1, y_2, ..., y_T\}$, is zero and its conditional MSE is given by

$$Var_T\left(y_{T+k}\right) = E_T\left(\sigma_{T+k}^2\right) = \frac{\omega}{1-\alpha-\beta} + (\alpha+\beta)^{k-1}\left(\sigma_{T+1}^2 - \frac{\omega}{1-\alpha-\beta}\right). \tag{4}$$

If ε_t is further assumed to be a Gaussian process then y_t is conditionally Gaussian, and one step ahead forecast errors are normally distributed. Therefore, 95% prediction intervals of y_{T+1} are given by $\pm 1.96\sigma_{T+1}$. However, the prediction error distribution when forecasting k-periods ahead for $k > 1$, is not normal even under the Gaussianity assumption. However, the usual approximation to the $(1-\alpha)\%$ prediction interval of returns for $k > 1$ is given by

$$\pm z_{\alpha/2} E_T\left(\sigma_{T+k}\right), \tag{5}$$

where $z_{\alpha/2}$ is the $\alpha/2$ quantile of the standard normal density.

Alternatively, Baillie and Bollerslev (1992) propose an improvement of the intervals in (5) based on Cornish-Fisher expansions making parametric assumptions on the distribution of ε_t. However, the prediction error distribution is only known for the GARCH(1,1) model and it seems difficult to generalize it to general GARCH(p,q) processes.

With respect to the prediction of future values of volatilities, the point predictor of σ^2_{T+k} is given by (4). Although, Baillie and Bollerslev (1992) derive the expression for the conditional MSE for the k-step ahead predictor of the conditional variance, the prediction error distribution for the conditional variance is not derived, and therefore, prediction intervals can not be obtained.

3. BOOTSTRAP PREDICTION INTERVALS

In this section we describe a bootstrap procedure that extends the procedure proposed in Pascual *et al.* (1998) for ARIMA models, to obtain prediction densities of future values of returns and volatilities of series generated by GARCH processes. For simplicity in the exposition and because the GARCH(1,1) is the model commonly used in practise, we concentrate on it hereafter. In this paper, we follow the mainstream of the literature in assuming that the specification of the model is known.

Let $\{y_1, \ldots, y_T\}$ be a sequence of T observations generated by a GARCH(1,1) process given by equation (1). The objective is to estimate directly the distribution of y_{T+k} and σ_{T+k} conditional on the available data. The unknown parameters (ω, α, β) are estimated by quasi-maximum likelihood (QML), maximizing the Gaussian log-likelihood function even if the assumption of normality is violated. Bollerslev and Woodridge (1992) prove that, under standard conditions, the QML estimator is consistent and has a limiting normal distribution.

Once the values (ω, α, β) are estimated by QML, say $\left(\widehat{\omega}, \widehat{\alpha}, \widehat{\beta}\right)$, the conditional variances are estimated by

$$\widehat{\sigma}_t^2 = \widehat{\omega} + \widehat{\alpha} y_{t-1}^2 + \widehat{\beta}\widehat{\sigma}_{t-1}^2, \; t = 2, ..., T, \tag{6}$$

with $\widehat{\sigma}_1^2 = \frac{\widehat{\omega}}{1-\widehat{\alpha}-\widehat{\beta}}$, the estimated marginal variance. Then, the residuals are computed by the standardized observations given by, $\widehat{\varepsilon}_t = y_t/\widehat{\sigma}_t$, $t = 1, ..., T$.

To incorporate the uncertainty due to parameter estimation, it is necessary to obtain bootstrap replicates $\{y_1^*, ..., y_T^*\}$ that mimic the structure of the original series. These replicates are obtained from the following recursion

$$\begin{aligned} \widehat{\sigma}_t^{*2} &= \widehat{\omega} + \widehat{\alpha} y_{t-1}^{*2} + \widehat{\beta}\widehat{\sigma}_{t-1}^{*2}, \\ y_t^* &= \varepsilon_t^* \widehat{\sigma}_t^*, \text{ for } t = 1, ..., T, \end{aligned} \tag{7}$$

where ε_t^* are random draws from $\widehat{F}_T$, the empirical distribution function of the centered residuals and the initial value for the volatility is given by $\sigma_1^{*2} = \widehat{\sigma}_1^2$. Once the parameters of this bootstrap series are estimated by QML, say $\left(\widehat{\omega}^*, \widehat{\alpha}^*, \widehat{\beta}^*\right)$, bootstrap forecasts of future values are obtained through the following recursions:

$$\begin{aligned} \widehat{\sigma}_{T+k}^{*2} &= \widehat{\omega}^* + \widehat{\alpha}^* y_{T+k-1}^{*2} + \widehat{\beta}^* \widehat{\sigma}_{T+k-1}^{*2}, \\ y_{T+k}^* &= \varepsilon_{T+k}^* \widehat{\sigma}_{T+k}^*, \text{ for } k = 1, 2, ... \end{aligned} \tag{8}$$

with ε_{T+k}^* being random draws from $\widehat{F}_T$ and the initial values are given by $y_T^* = y_T$, and

$$\widehat{\sigma}_T^{*2} = \frac{\widehat{\omega}^*}{1 - \widehat{\alpha}^* - \widehat{\beta}^*} + \widehat{\alpha}^* \sum_{j=0}^{T-2} \widehat{\beta}^{*j} (y_{T-j-1}^2 - \frac{\widehat{\omega}^*}{1 - \widehat{\alpha}^* - \widehat{\beta}^*}). \tag{9}$$

Notice that in expression (9) although $\widehat{\sigma}_T^{*2}$ is different in all bootstrap replicates, its value is obtained using the corresponding bootstrap parameter estimates and always the original series in such a way that its value is small when the returns at the end of the sample period are small and big when they are big in absolute value. Consequently, $\widehat{\sigma}_T^{*2}$ only incorporates the variability due to parameter estimation and takes into account the state of the process when predictions are going to be made.

Once we obtain a set of B bootstrap replicates for y_{T+k}, say $\left(y_{T+k}^{*(1)}, ..., y_{T+k}^{*(B)}\right)$, the prediction limits are defined as the quantiles of the bootstrap distribution function of y^*_{T+k}. More specifically, if $G^*_Y(h) = \Pr\left(y^*_{T+k} \le h\right)$ is the distribution function of y^*_{T+k} and its Monte Carlo estimate is $G^*_{Y,B}(h) = \#(y^{*b}_{T+k} \le h)/B$, a $100\gamma\%$ prediction interval for y^*_{T+k} is given by

$$\left[L^*_{Y,B}(y), U^*_{Y,B}(y)\right] = \left[Q^*_{Y,B}\left(\frac{1-\gamma}{2}\right), Q^*_{Y,B}\left(\frac{1+\gamma}{2}\right)\right] \quad (10)$$

where $Q^*_{Y,B} = G^{*-1}_{Y,B}$.

As stated previously, we can also obtain, at the same time, prediction intervals for the volatility k periods into the future. Given a set of B bootstrap replicates of the volatility for any horizon k, say $\left(\widehat{\sigma}^{*(1)}_{T+k}, ..., \widehat{\sigma}^{*(B)}_{T+k}\right)$ we proceed as before, using as prediction limits the quantiles of the bootstrap distribution function of $\widehat{\sigma}^*_{T+k}$. In this case, if $G^*_\sigma(h) = \Pr\left(\widehat{\sigma}^*_{T+k} \le h\right)$ is the distribution function of $\widehat{\sigma}^*_{T+k}$ and its Monte Carlo estimate is $G^*_{\sigma,B}(h) = \#(\widehat{\sigma}^{*b}_{T+k} \le h)/B$, a $100\gamma\%$ prediction interval for $\widehat{\sigma}^*_{T+k}$ is given by

$$\left[L^*_{\sigma,B}(y), U^*_{\sigma,B}(y)\right] = \left[Q^*_{\sigma,B}\left(\frac{1-\gamma}{2}\right), Q^*_{\sigma,B}\left(\frac{1+\gamma}{2}\right)\right] \quad (11)$$

where $Q^*_{\sigma,B} = G^{*-1}_{\sigma,B}$.

Summarizing, the steps for obtaining bootstrap prediction intervals are:

Step 1. Compute the centered residuals $\widehat{\varepsilon}_t$, and let $\widehat{F}_T$ denote their empirical distribution function.

Step 2. Generate a bootstrap series using the recursion in (7) and calculate the corresponding bootstrap estimates $\left(\widehat{\omega}^*, \widehat{\alpha}^*, \widehat{\beta}^*\right)$.

Step 3. Obtain bootstrap future values of returns and volatilities for any horizon k by the recursion in (8).

Step 4. Repeat the last two steps B times and then go to Step 5.

Step 5. The endpoints of the prediction intervals are given by quantiles of $G^*_{Y,B}$ and $G^*_{\sigma,B}$, the bootstrap distribution functions of y^*_{T+k} and $\widehat{\sigma}^*_{T+k}$, respectively.

Alternatively, the bootstrap procedure just described could be also applied to construct prediction intervals conditional on the parameter estimates; hereafter CB (conditional bootstrap). This procedure has been previously proposed by Miguel and Olave (1999) although they only focus on the construction of prediction intervals for y_{T+k} and do not consider the construction of prediction intervals for future volatilities. In this case, the estimated parameters are kept fixed in all bootstrap forecasts of y^*_{T+k} and $\widehat{\sigma}^*_{T+k}$ for k=1,2,.... Therefore, it is not necessary to generate bootstrap replicates of the series as in (7) and the bootstrap forecasts k-steps ahead depend only on the resampled residuals. The recursive equations of the CB are

$$\begin{aligned} \widehat{\sigma}^{*2}_{T+k} &= \widehat{\omega} + \widehat{\alpha} y^{*2}_{T+k-1} + \widehat{\beta}\widehat{\sigma}^{*2}_{T+k-1}, \\ y^*_{T+k} &= \varepsilon^*_{T+k}\widehat{\sigma}^*_{T+k}, \text{ for } k = 1, 2, ... \end{aligned} \tag{12}$$

where $y^*_T = y_T$ and $\widehat{\sigma}^{*2}_T = \widehat{\sigma}^2_T$ obtained using (6).

Since the parameter estimates are kept fixed in all bootstrap replicates of future values, the CB prediction intervals do not incorporate the uncertainty due to parameter estimation. Notice that $\widehat{\sigma}^{*2}_{T+1} = \widehat{\omega} + \widehat{\alpha} y^2_T + \widehat{\beta}\widehat{\sigma}^2_T$ is also kept fixed in all bootstrap replicates. Consequently, CB does not allow to estimate the one step ahead distribution of the volatility process. This could be expected as in GARCH models, the variance is observable one period ahead.

Notice that, although we have described the construction of prediction intervals for GARCH(1,1) models, the proposed bootstrap method can be easily generalized to deal with general GARCH(p,q) processes.

4. ASYMPTOTIC PROPERTIES OF BOOTSTRAP PREDICTION INTERVALS

In this section, we analyze the asymptotic properties of the proposed bootstrap procedure. Here and elsewhere $O_p(1)$ stands for boundedness in probability, while $o_p(1)$ denotes convergence to zero in probability.

Let $Y^*_T \equiv \{y^*_1, ..., y^*_T\}$ be a bootstrap sample generated following the resampling scheme described in the previous section and let P^* denote the underlying probability measure. We assume that $\widehat{\theta}_T^* = \left(\widehat{\omega}^*, \widehat{\alpha}^*, \widehat{\beta}^*\right)$ is the corresponding QML estimator calculated at the bootstrap sample, that is, is a function of the data Y_T^* satisfying the first order conditions,

$$\frac{\partial L_T(\theta)}{\partial \theta} \Big|_{\theta=\widehat{\theta}_T^*} = 0. \tag{13}$$

Let $\theta_0 = (\omega_0, \alpha_0, \beta_0)'$ be the true parameter vector, then Lumsdaine (1996) proves that θ_o is the unique maximizer of the Gaussian log-likelihood function $L(\theta) = \lim_{T\to\infty} L_T(\theta)$, and as $T \to \infty$,

$$\widehat{\theta}_T - \theta_o \to 0 \text{ in probability,}$$

and

$$T^{1/2}\left(\widehat{\theta}_T - \theta_o\right) \xrightarrow{d} N\left(0, A_o^{-1} B_o A_o^{-1}\right),$$

where

$$A_o = -E\left[\frac{\partial^2 l_t(\theta_o)}{\partial\theta\partial\theta'}\right], \text{ and } B_o = TE\left[\frac{\partial L_T(\theta_o)}{\partial\theta}\frac{\partial L_T(\theta_o)}{\partial\theta'}\right],$$

under the following two assumptions.

Assumption 1: The true parameter vector $\theta_o \in \Theta \subseteq \mathbb{R}^3$ is in the interior of Θ, a compact, convex parameter space. Specifically, for any vector $(\omega, \alpha, \beta) \in \Theta$, assume that $\delta \leq \omega \leq W$, $\delta \leq \alpha \leq (1-\delta)$ and $\delta \leq \beta \leq (1-\delta)$ for some constant $\delta > 0$, where W and δ are given a priori, and $E\left[\ln\left(\alpha\varepsilon_t^2 + \beta\right)\right] < 0$.

Assumption 2: $\{\varepsilon_t\}$ is i.i.d., drawn from a symmetric, unimodal density, bounded in a neighborhood of 0, with mean 0, variance 1, and $E\left(\varepsilon_t^{32}\right) < \infty$. In addition, σ_t^2 is independent of $\{\varepsilon_t, \varepsilon_{t+1}, ...\}$.

Defining the following

$$\frac{1}{\gamma} \equiv \left\{E\left[\left(\alpha_0\varepsilon_t^2 + \beta_0\right)^{-8}\right]\right\}^{\frac{1}{8}},$$

we can write $\gamma = \beta_0 + R_1 + R_2$, for some constants $R_1 > 0$ and $R_2 > 0$. Then, define Θ^+ to be the restriction of the parameter space Θ to $0 < \delta \leq \omega \leq W$, $0 < \delta \leq \alpha \leq (1-\delta)$

and $0 < \delta \leq \beta \leq \gamma - R_2 < \gamma$. All the results proven in Lumsdaine (1996) refers to this subset of Θ, which defines a local (nonzero measure) neighborhood of the true parameter space.

To establish the asymptotic justification of the bootstrap procedure, Assumptions 1 and 2 should be verified for each bootstrap sample. Moreover, we will prove that these Assumptions are satisfied in conditional probability. To achieve this objective, the following two Lemmas are needed.

Lemma 1 *Under the usual stationary assumptions for the GARCH(1,1) model, it follows that* $\widehat{\sigma}_t^2 - \sigma_t^2 = o_p(1) + \beta^t O_p(1)$.

Proof. Recall that for any given t, the conditional variance in (2) is given by

$$\sigma_t^2(\theta) = \frac{\omega}{1-\alpha-\beta} + \alpha \sum_{j=0}^{t-2} \beta^j z_{t-j-1},$$

where $z_t = y_t^2 - \frac{\omega}{1-\alpha-\beta}$ and denote by $\widehat{\sigma}_t^2 = \sigma_t^2\left(\widehat{\theta}_T\right)$and $\widehat{z}_t = y_t^2 - \frac{\widehat{\omega}}{1-\widehat{\alpha}-\widehat{\beta}}$. In this case,

$$\begin{aligned}
\widehat{\sigma}_t^2 - \sigma_t^2 &= \left(\frac{\widehat{\omega}}{1-\widehat{\alpha}-\widehat{\beta}} - \frac{\omega}{1-\alpha-\beta}\right) + \sum_{j=0}^{t-2}\left(\widehat{\alpha\beta^j}\,\widehat{z}_{t-j-1} - \alpha\beta^j z_{t-j-1}\right) \\
&\quad -\alpha \sum_{j=t-1}^{\infty} \beta^j z_{t-j-1} \\
&= \left(\frac{\widehat{\omega}}{1-\widehat{\alpha}-\widehat{\beta}} - \frac{\omega}{1-\alpha-\beta}\right) + \sum_{j=0}^{t-2}\left(\widehat{\alpha\beta^j} - \alpha\beta^j\right) y_{t-j-1}^2 \\
&\quad + \sum_{j=0}^{t-2}\left\{\widehat{\alpha\beta^j}\left(\frac{\widehat{\omega}}{1-\widehat{\alpha}-\widehat{\beta}}\right) - \alpha\beta^j\left(\frac{\omega}{1-\alpha-\beta}\right)\right\} - \alpha \sum_{j=t-1}^{\infty} \beta^j z_{t-j-1}.
\end{aligned}$$

The first three terms go to zero in probability because of the $\sqrt{T}$−consistency of $\left\{\widehat{\theta}_T\right\}$. For the last term, since $z_j = O_p(1)$, we have that $\alpha\sum_{j=t-1}^{\infty}\beta^j z_{t-j-1} = \alpha\sum_{j=t-1}^{\infty}\beta^j O_p(1) = \beta^t O_p(1)$ obtaining the result. ■

The next result states that the empirical distribution of the centered residuals, $\widehat{F}_T$, approximates the true distribution function F, with the help of the Mallows metric.

Lemma 2 *$d_2\left(F, \widehat{F}_T\right) \to 0$ in probability as $T \to \infty$.*

Proof. We will prove that the Mallows distance between the true distribution function of the errors, F, and the empirical distribution function of the centered residuals, $\widehat{F}_T$, goes to zero in probability as $T \to \infty$. Because d_2 is a metric

$$d_2(\widehat{F}_T, F)^2 \leq d_2(\widehat{F}_T, F_T)^2 + d_2(F_T, F)^2.$$

From Bickel and Freedman (1981) we have that $d_2(F_T, F) \to 0$ as $T \to \infty$ almost everywhere, hence, we just have to prove the consistence to zero of the other term.

Next, let J be Laplace distributed on $\{1, 2, \cdots, T\}$, i.e. J=j with probability 1/T for each j=1,$\cdots, T$, and define random variables X_1 and Y_1 with marginal F_T and $\widehat{F}_T$ respectively according to

$$X_1 = \varepsilon_J \ , \ Y_1 = \widehat{\varepsilon}_J - \widehat{\varepsilon}_{\cdot}.$$

Observe that,

$$\begin{aligned} d_2(\widehat{F}_T, F_T)^2 &= \inf\ E(X-Y)^2 \leq E(X_1 - Y_1)^2 \\ &= \frac{1}{T}\sum_{j=1}^{T}(\widehat{\varepsilon}_j - \varepsilon_j - \frac{1}{T}\sum_{i=1}^{T}\widehat{\varepsilon}_i)^2 \\ &\leq \frac{6}{T}\sum_{j=1}^{T}(\widehat{\varepsilon}_j - \varepsilon_j)^2 + \frac{3}{T^2}(\sum_{i=1}^{T}\varepsilon_i)^2 \\ &= \frac{6}{T}\sum_{j=1}^{T}\frac{y_j^2}{\widehat{\sigma}_j^2\sigma_j^2}(\widehat{\sigma}_j - \sigma_j)^2 + \frac{3}{T^2}(\sum_{i=1}^{T}\varepsilon_i)^2 \end{aligned}$$

Using the usual central limit theorem, $\frac{1}{\sqrt{T}}\sum_{j=1}^{T}\varepsilon_j = O_p(1)$, we have the convergency to zero in probability of the second term. For the first term, we have that $\frac{y_j^2}{\widehat{\sigma}_j^2\sigma_j^2} = O_p(1)$, and $(\widehat{\sigma}_j - \sigma_j)^2 = o_p(1) + \beta^j O_p(1)$. To prove this last assertion, note that $\widehat{\sigma}_j^2 - \sigma_j^2 = (\widehat{\sigma}_j - \sigma_j)(\widehat{\sigma}_j + \sigma_j) = o_p(1) + \beta^j O_p(1)$ by Lemma 1. Therefore, because $(\widehat{\sigma}_j + \sigma_j) = O_p(1)$, we have that $(\widehat{\sigma}_j - \sigma_j) = o_p(1) + \beta^j O_p(1)$. Consequently,

$(\widehat{\sigma}_j - \sigma_j)^2 = o_p(1) + \beta^j O_p(1)$. At this point,

$$\begin{aligned}
\frac{6}{T}\sum_{j=1}^{T}\frac{y_j^2}{\widehat{\sigma}_j^2\sigma_j^2}(\widehat{\sigma}_j - \sigma_j)^2 &= \frac{6}{T}\sum_{j=1}^{T} O_p(1)\left[o_p(1) + \beta^j O_p(1)\right] \\
&= \frac{6}{T}\sum_{j=1}^{T} O_p(1) o_p(1) + \frac{6}{T}\sum_{j=1}^{T}\beta^j O_p(1) \\
&= \frac{6}{T}\sum_{j=1}^{T} o_p(1) + \frac{6\beta}{1-\beta}\frac{(1-\beta^T)}{T} O_p(1),
\end{aligned}$$

and the first term of this last expression goes to zero in probability as T goes to infinity and the same for the other term because $0 \leq \beta < 1$, which concludes the result. ■

Now, we are ready to justify the bootstrap procedure verifying Assumptions 1 and 2 for each bootstrap sample in conditional probability. For that, we give the bootstrap version for these two assumptions and prove both of them.

Assumption 1*: The parameter vector $\widehat{\theta}_T \in \Theta \subseteq \mathbb{R}^3$ is in the interior of Θ (*in probability*), a compact, convex parameter space. Specifically, for any vector $(\omega, \alpha, \beta) \in \Theta$, assume that $\delta \leq \omega \leq W$, $\delta \leq \alpha \leq (1-\delta)$ and $\delta \leq \beta \leq (1-\delta)$ for some constant $\delta > 0$, where W and δ are given a priori, and $E^*\left[\ln\left(\alpha\varepsilon_t^{*2} + \beta\right)\right] < 0$ (*in probability*).

Assumption 2*: $\{\varepsilon_t^*\}$ is i.i.d., drawn from a symmetric, unimodal density, bounded in a neighborhood of 0, with mean 0, variance 1 (*in probability*), and $E^*\left(\varepsilon_t^{*32}\right) < \infty$. In addition, $\widehat{\sigma}_t^2$ is independent of $\left\{\varepsilon_t^*, \varepsilon_{t+1}^*, \ldots\right\}$ (*in probability*).

For verifying Assumptions 1* and 2*, we need to assume the following hypotheses..

Hypotheses 1. $\sup_T E\left[\left|\ln\left(\alpha\widehat{\varepsilon}_t^2 + \beta\right)\right|^{2+\varsigma}\right] < \infty$ for some $\varsigma > 0$.

Hypotheses 2. $\sup_T E\left[\left|\widehat{\varepsilon}_t\right|^{2+\varsigma}\right] < \infty$ for some $\varsigma > 0$.

Since θ_o is in the interior of Θ, we can find $\eta > 0$ such that $B(\theta_o, \eta) \subset \overset{\circ}{\Theta}$. Even more, by theorem 2 in Lumsdaine (1996) we have that $\widehat{\theta}_T \to \theta_0$ in probability as $T \to \infty$. Therefore, given $\eta/2$, exists $n_o \in \mathbb{N}$ such that, for all $n \geq n_o$, $\widehat{\theta}_T \in B(\theta_o, \eta/2) \subset$

$B(\theta_o, \eta)$, concluding that $\widehat{\theta}_T \in \Theta \subseteq \mathbb{R}^3$ is in the interior of Θ (*in probability*).

With respect to $E^*\left[\ln\left(\alpha\varepsilon_t^{*2} + \beta\right)\right] < 0$ (*in probability*), note that

$$E^*\left[\ln\left(\alpha\varepsilon_t^{*2} + \beta\right)\right] = \frac{1}{T}\sum_{j=1}^{T} \ln\left(\alpha\widehat{\varepsilon}_j^2 + \beta\right).$$

By the Weak Law of Large Numbers, the last sum converges in probability to $E\left[\ln\left(\alpha\widehat{\varepsilon}_j^2 + \beta\right)\right]$. At the same time, we have by Lemma 2 that $\ln\left(\alpha\widehat{\varepsilon}_j^2 + \beta\right) \xrightarrow{d} \ln\left(\alpha\varepsilon_j^2 + \beta\right)$ in probability. Now, by Hypotheses 1, we have that the sequence $\left\{\ln\left(\alpha\widehat{\varepsilon}_t^2 + \beta\right)\right\}$ is uniformly integrable and then, by Theorem 25.12 in Billingsley (1986) we have that

$$E\left[\ln\left(\alpha\widehat{\varepsilon}_j^2 + \beta\right)\right] \to E\left[\ln\left(\alpha\varepsilon_j^2 + \beta\right)\right] \text{ in probatility.}$$

These two convergencies written together, prove that,

$$E^*\left[\ln\left(\alpha\varepsilon_t^{*2} + \beta\right)\right] < 0 \ (\textit{in probability}).$$

Using the same kind of arguments it is easy to verify Assumption 2*. First of all, we have that the sequence of $\{\varepsilon_t^*\}$ is i.i.d from $\widehat{F}_T$, the empirical distribution function of the centered residuals; therefore $E^*[\varepsilon_t^*] = 0$; To see that $Var^*[\varepsilon_t^*] = 1$ (*in probability*), just note that

$$\begin{aligned} Var^*[\varepsilon_t^*] &= E^*\left[\varepsilon_t^{*2}\right] - \left\{E^*[\varepsilon_t^*]\right\}^2 \\ &= \frac{1}{T}\sum_{j=1}^{T}\widehat{\varepsilon}_t^2 - \left\{\frac{1}{T}\sum_{j=1}^{T}\widehat{\varepsilon}_t\right\}^2 \\ &= \frac{1}{T}\sum_{j=1}^{T}\widehat{\varepsilon}_t^2 \xrightarrow{p} E\left[\widehat{\varepsilon}_t^2\right]. \end{aligned}$$

Now, by hypothesis 2, $\left\{\widehat{\varepsilon}_t^2\right\}$ is uniformly integrable, then, $E\left[\widehat{\varepsilon}_t^2\right] \to 1$ in probability, obtaining the result. Even more, $E^*\left(\varepsilon_t^{*32}\right) < \infty$, since $E^*\left(\varepsilon_t^{*32}\right) = \frac{1}{T}\sum_{j=1}^{T}\widehat{\varepsilon}_t^{32}$ which is well defined.

Finally, to prove that $\widehat{\sigma}_t^{*2}$ is independent of $\left\{\varepsilon_t^*, \varepsilon_{t+1}^*, ...\right\}$ (*in probability*) just note that

$$\begin{aligned}
\widehat{\sigma}_t^{*2} &= \widehat{\omega} + \widehat{\alpha} y_{t-1}^{*2} + \widehat{\beta}\widehat{\sigma}_{t-1}^{*2} \\
&= \omega + \alpha y_{t-1}^{*2} + \beta \widehat{\sigma}_{t-1}^{*2} + o_p(1) \\
&= f\left(\omega, \alpha, \beta, \varepsilon_{t-1}^*, ..., \varepsilon_1^*\right) + o_p(1),
\end{aligned}$$

and then, since the sequence $\{\varepsilon_t^*\}$ is i.i.d. we have the result.

Once Assumptions 1* and 2* have been checked, we have available all the results in Lumsdaine (1996) for any bootstrap sample (in probability), and in particular, we ensure that $\sqrt{T}\left(\widehat{\theta}_T^* - \widehat{\theta}_T\right) = O_{P^*}(1)$ in probability.

We conclude that $\widehat{\theta}_T^* - \theta_0 \xrightarrow{P^*} 0$ in probability as T goes to ∞, for θ_0 the true parameter value. This is achieved by using the usual triangular inequality.

At this point, we are ready to proof that bootstrap future returns and volatilities converge in conditional distribution (in probability) to the corresponding true variables as the sample size increases.

Lemma 3 *Let* $\{y_{T-n+1}, \ldots, y_T\}$ *be a realization of size* n *of a stationary GARCH(1,1) process defined by* (1)*. Then, it follows that* $\widehat{\sigma}_T^{*2} - \sigma_T^2 = o_{p^*}(1) + \beta^n O_{p^*}(1)$ *in probability.*

Proof. Using the same notation as in Lemma 1 and denoting by $\widehat{\sigma}_T^{*2} = \sigma_T^2\left(\widehat{\theta}_n^*\right)$, it is possible to obtain the following expression

$$\begin{aligned}
\widehat{\sigma}_T^{*2} - \sigma_T^2 &= \left(\frac{\widehat{\omega}^*}{1 - \widehat{\alpha}^* - \widehat{\beta}^*} - \frac{\omega}{1 - \alpha - \beta}\right) + \sum_{j=0}^{n-2}\left(\widehat{\alpha}^*\widehat{\beta}^{*j} - \alpha\beta^j\right) y_{T-j-1}^2 \\
&+ \sum_{j=0}^{n-2}\left\{\widehat{\alpha}^*\widehat{\beta}^{*j}\left(\frac{\widehat{\omega}^*}{1 - \widehat{\alpha}^* - \widehat{\beta}^*}\right) - \alpha\beta^j\left(\frac{\omega}{1 - \alpha - \beta}\right)\right\} - \alpha\sum_{j=n-1}^{\infty}\beta^j z_{T-j-1}.
\end{aligned}$$

The first three terms goes to zero in conditional probability (in probability) as stated previously.. For the last term, since $z_j = O_p(1)$, $\alpha\sum_{j=n-1}^{\infty}\beta^j z_{t-j-1} = \beta^n O_p(1)$ obtaining the result. ∎

Theorem 4 *Let* $\{y_{T-n+1},\ldots,y_T\}$ *be a realization of size* n *of a stationary GARCH(1,1) process given in* (1)*. Let* $\left(\widehat{\omega},\widehat{\alpha},\widehat{\beta}\right)$ *be any QML-estimate of* (ω,α,β) *and let* Y^*_{T+k} *be obtained following steps 1 to 5. Then, for any distance d metrizing weak convergence,* $d\left(Y^*_{T+k},Y_{T+k}\right)\to 0$ *in probability as* $n\to\infty$*.*

Proof. We achieve the result using back-substitution. For k=1, the result is trivial. In this case,

$$y^*_{T+1}=\varepsilon^*_{T+1}\left(\widehat{\omega}^*+\widehat{\alpha}^*y_T^2+\widehat{\beta}^*\widehat{\sigma}_T^{*2}\right)^{1/2}.$$

By Lemma 2, $\varepsilon^*_{T+1}\to\varepsilon_{T+1}$ in conditional distribution (in probability) as n goes to ∞. Even more, by Lemmas 3 and 4, $\left(\widehat{\omega}^*,\widehat{\alpha}^*,\widehat{\beta}^*,\widehat{\sigma}_T^{*2}\right)\to(\omega,\alpha,\beta,\sigma_T^2)$ in conditional probability (in probability) as n goes to ∞.

Let $a=(a_0,a_1,a_2,a_3)'$ be a vector in $\mathbb{R}^4$ where the first three elements fulfill the usual stationarity conditions for the GARCH(1,1) model (see section 2), and for any horizon k let define the vector $X(k)=(x_1,\ldots,x_k)'$ in $\mathbb{R}^k$. Then, the function

$$g_1(a,X(1))=x_1\left(a_0+a_1y_T^2+a_2a_3^2\right)^{1/2}$$

is continuous at $(a,X(1))$ for all $X(1)$ in $\mathbb{R}$.

In such a case, $y^*_{T+1}=g_1\left(\widehat{b}^*_n,\varepsilon^*(1)\right)$ and $y_{T+1}=g_1(b,\varepsilon(1))$, for $\widehat{b}^*_n=\left(\widehat{\omega}^*,\widehat{\alpha}^*,\widehat{\beta}^*,\widehat{\sigma}_T^{*2}\right)'$, $b=(\omega,\alpha,\beta,\sigma_T^2)'$, $\varepsilon^*(1)=\left(\varepsilon^*_{T+1}\right)$ and $\varepsilon(1)=(\varepsilon_{T+1})$. Then, by the bootstrap version of Slutzky's theorem,

$$g_1\left(\widehat{b}^*_n,\varepsilon^*(1)\right)\to g_1(b,\varepsilon(1))\ \text{ in conditional distribution}$$

(in probability), as $n\to\infty$, since $\widehat{b}^*_n\to b$ in conditional probability (in probability) and $\varepsilon^*(1)\to\varepsilon(1)$ in conditional distribution (in probability).

For forecast horizon k=2, using a recursive argument, we express y^*_{T+2} as

$$y^*_{T+2}=\varepsilon^*_{T+2}\left(\widehat{\omega}^*\left(1+\widehat{\beta}^*\right)+\widehat{\alpha}^*g_1^2\left(\widehat{b}^*_n,\varepsilon^*(1)\right)+\widehat{\alpha}^*\widehat{\beta}^*y_T^2+\widehat{\beta}^{*2}\widehat{\sigma}_T^{*2}\right)^{1/2}.$$

The function

$$g_2(a, X(2)) = x_2 \left(a_0 (1 + a_2) + a_1 g_1^2 (a, X(1)) + a_1 a_2 y_T^2 + a_2^2 a_3^2\right)^{1/2},$$

is continuous at $(a, X(2))$ for all $X(2)$ in $\mathbb{R}^2$, and is such that $y_{T+2}^* = g_2\left(\widehat{b}_n^*, \varepsilon^*(2)\right)$ and $y_{T+2} = g_2(b, \varepsilon(2))$, with $\widehat{b}_n^*$ and b as before, and $\varepsilon^*(2) = \left(\varepsilon_{T+1}^*, \varepsilon_{T+2}^*\right)$ and $\varepsilon(2) = (\varepsilon_{T+1}, \varepsilon_{T+2})$. Since ε_{T+1}^* and ε_{T+2}^* are independent random draws, we have that $\varepsilon^*(2) \to \varepsilon(2)$ in conditional distribution (in probability) by Cramer-Wald theorem. Therefore, by Slutzky's theorem,

$$g_2\left(\widehat{b}_n^*, \varepsilon^*(2)\right) \to g_2(b, \varepsilon(2)) \text{ in conditional distribution}$$

(in probability), as $n \to \infty$, that is,

$$y_{T+2}^* \to y_{T+2} \text{ in conditional distribution}$$

(in probability), as $n \to \infty$.

For a general horizon k, by the recursive argument, we express y_{T+k}^* as

$$y_{T+k}^* = \varepsilon_{T+k}^* \left(\widehat{\omega}^* \sum_{j=0}^{k-1} \widehat{\beta}^{*j} + \widehat{\alpha}^* \widehat{\beta}^{*k-1} y_T^2 + \widehat{\beta}^{*k} \widehat{\sigma}_T^{*2} + \widehat{\alpha}^* \sum_{j=0}^{k-1} \widehat{\beta}^{*k-1-j} g_j^2 \left(\widehat{b}_n^*, \varepsilon^*(j)\right)\right)^{1/2}.$$

Then, the function

$$g_k(a, X(k)) = x_k \left(a_0 \sum_{j=0}^{k-1} a_2^j + a_1 a_2^{k-1} y_T^2 + a_2^k a_3^2 + a_1 \sum_{j=0}^{k-1} a_2^{k-1-j} g_j^2 (a, X(j))\right)^{1/2}$$

is continuous at $(a, X(k))$ for all $X(k)$ in $\mathbb{R}^k$ since, by definition, g_k is the composition of continuous functions g_j for j=1,2,...,k-1. In such a case, $y_{T+k}^* = g_k\left(\widehat{b}_n^*, \varepsilon^*(k)\right)$ and $y_{T+k} = g_k(b, \varepsilon(k))$, where $\varepsilon^*(k) = \left(\varepsilon_{T+1}^*, \ldots, \varepsilon_{T+k}^*\right)$ and $\varepsilon(k) = (\varepsilon_{T+1}, \ldots, \varepsilon_{T+k})$. Since $\varepsilon_{T+1}^*, \ldots, \varepsilon_{T+k}^*$ are independent random draws, we have that $\varepsilon^*(k) \to \varepsilon(k)$ in conditional distribution (in probability) by Cramer-Wald´s theorem. Therefore, by Slutzky´s theorem,

$$g_k\left(\widehat{b}_n^*, \varepsilon^*(k)\right) \to g_k(b, \varepsilon(k)) \text{ in conditional distribution}$$

(in probability), as $n \to \infty$, that is,

$$y^*_{T+k} \to y_{T+k} \text{ in conditional distribution}$$

(in probability), as $n \to \infty$. ■

Theorem 5 *Let $\{y_{T-n+1}, \ldots, y_T\}$ be a realization of size* n *of a stationary GARCH(1,1) process given in* (1). *Let $\left(\widehat{\omega}, \widehat{\alpha}, \widehat{\beta}\right)$ be any QML-estimate of (ω, α, β) and let σ^*_{T+k} be obtained following steps 1 to 5. Then, for any distance d metrizing weak convergence, $d\left(\sigma^*_{T+k}, \sigma_{T+k}\right) \to 0$ in probability as $n \to \infty$.*

Proof. We follow the same strategy as in theorem 5. For k=1 the result is obvious since in this case,

$$\sigma^{*2}_{T+1} = \widehat{\omega}^* + \widehat{\alpha}^* y_T^2 + \widehat{\beta}^* \widehat{\sigma}^{*2}_T$$

and then, $\sigma^{*2}_{T+1} \to \sigma^2_{T+1}$ in conditional probability (in probability) as n goes to ∞ by Lemmas 3 and 4.

For forecast horizon k=2, we can express σ^{*2}_{T+2} as

$$\sigma^{*2}_{T+2} = \widehat{\omega}^* \left(1 + \widehat{\beta}^*\right) + \widehat{\alpha}^* g_1^2 \left(\widehat{b}^*_n, \varepsilon^*(1)\right) + \widehat{\alpha}^* \widehat{\beta}^* y_T^2 + \widehat{\beta}^{*2} \widehat{\sigma}^{*2}_T$$

where $g_1(a, X(1))$ is a continuous function defined in theorem 5. Now, using the same notation as in that theorem, we define the continuous function at $(a, X(1))$ for all $X(1)$ in $\mathbb{R}$,

$$h_2(a, X(1)) = a_0(1 + a_2) + a_1 g_1^2(a, X(1)) + a_1 a_2 y_T^2 + a_2^2 a_3^2.$$

By construction, this function is such that $\sigma^{*2}_{T+2} = h_2\left(\widehat{b}^*_n, \varepsilon^*(1)\right)$ and $\sigma^2_{T+2} = h_2(b, \varepsilon(1))$ and then, using the same arguments as in theorem 5, by Slutzky´s theorem $\sigma^{*2}_{T+2} \to \sigma^2_{T+2}$ in conditional distribution (in probability) as n goes to ∞.

For a general horizon k we express σ^{*2}_{T+k} as

$$\sigma^{*2}_{T+k} = \widehat{\omega}^* \sum_{j=0}^{k-1} \widehat{\beta}^{*j} + \widehat{\alpha}^* \widehat{\beta}^{*k-1} y_T^2 + \widehat{\beta}^{*k} \widehat{\sigma}^{*2}_T + \widehat{\alpha}^* \sum_{j=0}^{k-1} \widehat{\beta}^{*k-1-j} g_j^2 \left(\widehat{b}^*_n, \varepsilon^*(j)\right).$$

Then, the function

$$h_k\left(a, X\left(k-1\right)\right) = a_0 \sum_{j=0}^{k-1} a_2^j + a_1 a_2^{k-1} y_T^2 + a_2^k a_3^2 + a_1 \sum_{j=0}^{k-1} a_2^{k-1-j} g_j^2\left(a, X\left(j\right)\right)$$

is continuous at $\left(a, X\left(k-1\right)\right)$ for all $X\left(k-1\right)$ in $\mathbb{R}^{k-1}$ since, by definition, h_k is the composition of continuous functions g_j for j=1,2,...,k-1. In such a case, $\sigma_{T+k}^{*2} = h_k\left(\widehat{b}_n^*, \varepsilon^*\left(k-1\right)\right)$ and $\sigma_{T+k}^2 = h_k\left(b, \varepsilon\left(k-1\right)\right)$. Therefore, by Slutzky's theorem,

$$h_k\left(\widehat{b}_n^*, \varepsilon^*\left(k-1\right)\right) \rightarrow h_k\left(b, \varepsilon\left(k-1\right)\right) \text{ in conditional distribution}$$

(in probability), as $n \rightarrow \infty$, that is,

$$\sigma_{T+k}^{*2} \rightarrow \sigma_{T+k}^2 \text{ in conditional distribution}$$

(in probability), as $n \rightarrow \infty$.

At this point, for any forecast horizon k, we also have the result for the volatilities, that is,

$$\sigma_{T+k}^{*} \rightarrow \sigma_{T+k} \text{ in conditional distribution}$$

(in probability), as $n \rightarrow \infty$, since the squared root is a continuous function in the range of $\left(b, \varepsilon\left(k-1\right)\right)$by hypotheses. ■

5. MONTE CARLO EXPERIMENTS

In this section, the finite sample behavior of the proposed bootstrap procedure to estimate prediction densities and intervals of future returns and volatilities of series generated by GARCH models is analyzed by means of Monte Carlo experiments. Prediction intervals of returns build by the proposed procedure (PRR) are compared with intervals based on the Normal approximation (STD) in (5) and with CB intervals.

We generate series by the following GARCH(1,1) model:

$$\begin{aligned} y_t &= \varepsilon_t \sigma_t \\ \sigma_t^2 &= 0.05 + 0.1 y_{t-1}^2 + 0.85 \sigma_{t-1}^2 \end{aligned} \tag{14}$$

with Gaussian, Student-t with 5 degrees of freedom, exponential centered to have zero mean and double-exponential innovations[1]. The sample sizes considered are 300, 1000 and 3000 observations. The prediction horizons under study are k=1, 2, 10 and 20 and the corresponding intervals are constructed with a nominal coverage 1-α equal to 0.80, 0.95 and 0.99. For each particular series generated by model (15) with a particular sample size and error distribution, we generate R=1000 future values of y_{T+k} and σ_{T+k} from that series and obtain 100α% prediction intervals for returns, denoted by (L_Y^*, U_Y^*) by each of the three procedures considered. PRR and CB intervals are constructed based on B=999 bootstrap replicates. The conditional coverage of each procedure is computed by

$$\widehat{\alpha}_Y^* = \# \left\{ L_Y^* \leq y_{T+k}^r \leq U_Y^* \right\} / R,$$

where y_{T+k}^r $(r = 1,\ldots,R)$ are future values of the variable generated previously.

At the same time we obtain a 100α% prediction interval for the volatility, denoted by (L_σ^*, U_σ^*) and estimate the conditional coverage of each procedure by

$$\widehat{\alpha}_\sigma^* = \# \left\{ L_\sigma^* \leq \sigma_{T+k}^r \leq U_\sigma^* \right\} / R,$$

where σ_{T+k}^r $(r = 1,\ldots,R)$ are the values of the future volatility generated previously.

The prediction intervals are compared in terms of average coverage and length, and the proportion of observations lying out to the left and to the right through all Monte Carlo replicates. The Monte Carlo results are based on 1000 replicates.

[1] Results for alternative models are similar to the ones reported and are avalaible from the authors upon request.

All computations have been carried out in a HP-UX C360 workstation, using Fortran 77 and the corresponding subroutines of Numerical Recipes by Press *et al.* (1986). In particular, Gaussian and Student-t errors are generated using the subroutine "*gasdev*" and the corresponding transformations in each case. Exponential errors are generated using uniform random numbers generated by subroutine "*rand2*" and transforming them as appropriate. The numerical optimization of the Gaussian log-Likelihood function has been carried out using the subroutine "*amoeba*" with the maximum allowed function evaluations set equal to 5000 and the fractional convergence tolerance set equal to 10^{-6}.

5.1. Prediction intervals for returns

Table 1 shows the results from the Monte Carlo experiments carried out with series generated by model (15) with ε_t being Gaussian. In this case, all the procedures considered to construct prediction intervals for returns have similar properties for all prediction horizons and nominal coverages. It is important to observe that although the conditional distribution of y_{T+k} is not normal, it seems that the normality approximation in (2) is not a bad assumption when building 95% prediction intervals. Also, notice that the performance of the bootstrap procedures is never worse than with the standard approach. Furthermore, when we compare PRR and CB intervals, small differences between them are observed. Therefore, introducing or not the variability due to the parameter estimation does not lead to an improvement in the performance of bootstrap prediction intervals for the returns series.

Tables 2 and 3 report the results for the same model with ε_t having a Student-t distribution with 5 degrees of freedom and for 80% and 99% prediction intervals respectively. In table 2 it is possible to observe that the average coverage and length of STD intervals in (5) is always over the empirical coverage even for very big sample sizes. On the other hand, in table 3 the STD average coverages and lengths are under the empirical values and quite different from the bootstrap values for the same quantities. As expected, the distortion does not disappear when the sample size increases. Consequently, as pointed out by Pascual, Romo and Ruiz (1998), standard intervals are clearly distorted when the error distribution is not Gaussian. On the other hand, comparing PRR and CB intervals, both have similar properties. It seems that for symmetric distributions, introducing the uncertainty due to parameter estimation in prediction intervals is not fundamental. Notice that the results in tables 2 and 3 are specially relevant in empirical applications. Bollerslev (1987) argues that the conditional normality assumption is not enough and that it is more adequate to use Student-t distributions. Therefore, the results in tables 2 and 3 show that in this case, it is more appropriate to predict using the PRR intervals. As an illustration, figure 1 represents kernel estimates of the empirical, the PRR and the standard normal densities for a particular series generated by model (15) with T=1000, for predictions made one-step-ahead. In figure 2, the same densities are plotted for predictions made twenty steps ahead. Notice that the PRR density is remarkably close to the empirical density while the standard normal is a worse approximation, not being able to represent the higher kurtosis in the data.

Finally, table 4 reports the results when the distribution of ε_t is exponential. In this case, STD intervals are clearly distorted. Comparing the resampling methods, we observe that for a sample size of 300, PRR clearly outperforms CB in the short term both in coverage and interval length, and the behavior tends to be similar as we go farther into the future. As expected, the differences between PRR and CB intervals disappear with the sample size. Figures 3 and 4 plot kernel estimates of the PRR densities together with the empirical and standard normal densities for a particular series generated by model (15) with exponential innovations and T=1000, for one-step and twenty-steps ahead respectively. As expected in this case, the standard density provides an inadequated approximation to the empirical distribution of futures returns. The PRR density is a better approximation although there is centered slightly to the left of the empirical distribution. In figures 5 and 6, where we represent the same densities as before with T=3000, it is possible to observe that the PRR distribution is a remarkable estimation of the empirical distribution.

5.2. Prediction intervals for volatilities

We now analyze the performance of PRR and CB prediction intervals for the future volatility itself. For this purpose, we use the same Monte Carlo design as for returns. In addition, we show the results for lead time k=2, since the behavior of the former differs with respect to k=1.

Table 5 reports the results for 95% prediction intervals when series are generated by model (15) with ε_t being Gaussian. Notice that in GARCH models the volatility is known one-step ahead so the only uncertainty associated with forecasting σ^2_{T+1} is due to parameter estimation. Therefore, the empirical length is not reported in table 5, since all the mass of the empirical distribution of σ_{T+1} conditional on the observed series is concentrated on σ_{T+1}. If the parameter estimates are considered as fixed the CB procedure is not able to give one-step ahead prediction intervals for volatilities. Notice that PRR intervals for future volatilities one-step ahead have average coverages close to the nominal values and that, as expected, their performance is better the bigger the sample size. Figures 7 and 8 plot kernel estimates of the PRR densities of one-step ahead volatilities of one series generated by model (15) with Gaussian innovations and T=1000 and 3000 respectively.

When forecasting two or more steps into the future, the average coverage of CB intervals is well under the nominal value. On the other hand, the average coverage of PRR intervals is closer to the nominal for all horizons considered. Although the average length of PRR intervals is over the empirical length for sample sizes of 300 observations, it gets closer as the sample size increases. Note that the empirical distribution of future values of volatilities is bounded by ω. Since, in practice, ω should be estimated, it is impossible to achieve exactly this bound with moderate sample sizes. In this case, the bootstrap estimate of the conditional distribution of σ^2_{T+k} is smoother than the empirical distribution, and tends to have values under this bound. This leads to give larger prediction intervals than the empirical ones, usually larger to the left, but with a good performance in terms of coverage. Figures 9 and 10 represent the empirical and PRR densities estimated for two-steps ahead prediction of volatilities generated by model (15) with Gaussian innovations and T=1000 and 3000 respectively. Notice that although for moderate sample sizes (T=1000) there are some distortions in the bootstrap density, when the sample size is big enough, the bootstrap density is able to represent the empirical density. In figures 11 and 12, we plot the same densities for predictions of the volatility twenty-steps ahead. For forecast horizons equal or greater than 2, the shape of the volatility is asymmetric and in concordance with the shapes usually found with real data; see for example Andersen et al. (1999).

Finally, table 6 reports results for the same model with ε_t having a Student-t distribution with 5 degrees of freedom. As in the Gaussian case, we may observe that the performance of the CB intervals is not adequate. Once more, the PRR intervals are too wide although the average length gets closer to the nominal length as the sample size increases.

Therefore, the results in tables 5 and 6 show the necessity of introducing the variability of the parameter estimates in order to obtain prediction intervals for the volatility with coverages close to the nominal values. When the variability of the estimated parameters is not introduced in the prediction intervals, the average coverage and average length are not adequate compared with the empirical values. Although when the sample size goes to infinity, CB is asymptotically correct, with the number of observations used in this study, it is still necessary to introduce the variability due to parameter estimation.

6. AN APPLICATION WITH REAL DATA

In this section we apply the bootstrap procedure described previously to construct prediction intervals for returns and volatilities of the Madrid Stock Market index IBEX-35. The estimation of the GARCH model used for the prediction of future returns and volatilities is based on daily closing prices of the IBEX-35 observed from 2/1/1996 to 3/3/2000, with a total of 1048 observations.

Daily percentage returns are obtained as first differences of logarithms scaled by multiplying by 100, i.e.,

$$R_t = 100 \log(P_t/P_{t-1}),$$

where P_t denotes the closing price at day t. The returns series is plotted in figure 13 where it is possible to observe volatility clustering with periods of low (high) volatility followed by periods of low (high) volatility suggesting the presence of conditional heteroscedasticity. Table 7 reports several sample moments of R_t. The estimated kurtosis coefficient is significantly bigger than 3 showing that the return distribution is leptokurtotic. Table 7 also contains the sample autocorrelations of returns and their squares. In this and subsequent tables, the standard deviations of sample autocorrelations of returns are corrected by ARCH effects as suggested by Diebold (1986) by the following expression

$$s.e.\,(r(k)) = \frac{1}{\sqrt{T}} \left\{ 1 + \frac{\widehat{\gamma}_2\,(k)}{[\widehat{\gamma}\,(0)]^2} \right\}^{\frac{1}{2}},$$

where $\widehat{\gamma}\,(0)$ is the sample variance of R_t and $\widehat{\gamma}_2\,(k)$ is the kth sample covariance of R_t^2. Notice that the order 1 autocorrelation of R_t is significative and that the squared returns are highly correlated.

Given that the order one autocorrelation is significative, the returns series has been corrected of this autocorrelation by fitting a MA(1) model with interventions with the following results:

$$R_t = \underset{(0.046)}{0.1188} + \widehat{a}_t + \underset{(0.047)}{0.1037}\,\widehat{a}_{t-1} - \underset{(0.273)}{5.2694}\,D_{1t} + \underset{(0.287)}{5.8184}\,D_{2t} - \underset{(0.134)}{6.9508}\,D_{3t} - \underset{(0.487)}{3.2550}\,D_{4t} \quad (15)$$

where D_{1t} is a pulse dummy variable that takes value 1 on the 4th of March of 1996 when the Partido Popular (PP) won the elections in Spain by a narrow margin and the Stock Market had a sharp decline. The second dummy variable, D_{2t}, corresponds to the 4th of January 1999 when the Euro was introduced. The last two dummy variables are due to extreme market reactions to external effects.

Table 8 reports the sample moments of the residuals from model (14). Once more, we observe excess kurtosis and significative autocorrelations of squares. From now on, the residuals from model (14) will be denoted by y_t.

To represent the dynamic evolution of squared residuals, we fit a GARCH(1,1) model to the filtered returns. The estimated model is

$$\widehat{\sigma}_t^2 = \underset{(0.011)}{0.0209} + \underset{(0.018)}{0.1060}\,y_{t-1}^2 + \underset{(0.020)}{0.8866}\,\widehat{\sigma}_{t-1}^2. \quad (16)$$

The sample moments of the standardized residuals, $\widehat{\varepsilon}_t = y_t/\widehat{\sigma}_t$, appear in table 9, where it can be observed that the GARCH model in (15) is able to represent adequately the dynamic structure of the volatility process. Although the excess kurtosis parameter is no longer significantly different from 3, the skewness coefficient significantly different from zero (p-value 0.04). The Jarque-Bera statistic for Normality equals 9.2996 (p-value=0.0095). Therefore, the standardized residuals are not Gaussian. Figure 14 represents a kernel estimate of the density of $\widehat{\varepsilon}_t$ together with the normal density.

Next, we apply the bootstrap procedure previously described to obtain prediction intervals of futures returns, y_{T+k}. The empirical out-of-sample forecast analysis is based on 20 temporal aggregates of tick by tick prices observed from 4/3/2000 to 31/3/2000. For each day in the period used to forecast, we have about 500 intraday one-minute observations. The estimated bootstrap densities for k=1 and 20, that correspond to predictions made one-day and one-month ahead, appear in Figure 15 where it is possible to observe the asymmetric shape previously observed in the standardized returns. Using these densities, we construct 80% and 95% PRR prediction intervals that have been plotted in figure 16 together with the point linear prediction of y_{T+k} (zero for all horizons), the observed returns and the corresponding prediction intervals constructed using the normal approximation. Once more we can see the asymmetry of the standardized returns, providing prediction intervals larger to the left than those obtained by standard methods.

Finally, we construct bootstrap prediction intervals for future volatilities. A common approach for judging prediction intervals is to check whether they contain the subsequent realizations of volatility with the desired coverage. However, as volatility is not directly observed, this approach is not immediately applicable for prediction interval evaluation. Different measures of volatility have been used in the literature to check whether GARCH models provide good forecasts of volatility. Andersen and Bollerslev (1998) propose to use a cumulative intraday squared-return measure of volatility. In this paper, we approximate the latent volatility at day t by the sum of the squared tick by tick returns during day t, i.e.

$$\sigma_t^2 = \sum R_{t,1}^2 + \cdots + R_{t,n}^2. \tag{17}$$

where n is the number of observations obtained at day t. Andersen (2000) points out that the magnitude of the measurement error of using σ_t^2 instead of the true volatility is inconsequential, and illustrates the use of the cumulative squared returns in (18) for volatility forecast evaluation.

In figure 17, we plot the histograms for the bootstrap predictions of volatilities 1, 2, 10 and 20 steps ahead into the future. Notice that the shape of volatility predictions is asymmetric and similar to the one obtained by Andersen et al. (1999) in the empirical analysis of very high frequency observations of exchange rates. Figure 18 shows point linear predictions of volatility together with the corresponding observed volatilities computed as in (18) and 80% and 95% PRR prediction intervals. In this figure we can see how the proposed resampling scheme gives good prediction intervals for both 80% and 95% in the sense that the 80% intervals leave 5 future values out when it is supposed to leave 4 and the 95% intervals leave 1 out that corresponds exactly with the nominal coverage.

7. CONCLUSIONS

In this paper, we have extended to GARCH processes the bootstrap procedure originally proposed by Pascual et al. (1998) to construct prediction intervals for ARIMA models. The bootstrap prediction intervals proposed can incorporated the uncertainty due to parameter estimation and do not rely on any assumption on the error distribution. Furthermore, incorporating the variability of the estimators, we can construct prediction intervals not only for future returns but also for volatilities.

We derive the asymptotic properties of the proposed bootstrap procedure and analyze its finite sample behavior by means of Monte Carlo experiments. The results of these experiments show that the standard prediction intervals for returns build treating the error distribution as if it were Normal for any prediction horizon, are adequate as far as the model is conditionally Normal. However, it has been often observed in empirical applications that the conditional distribution of the errors of GARCH models is leptokurtic; see, for example, Bollerslev et al. (1994) and the references therein. In this case, the standard prediction intervals for returns are not able to deal with non-Gaussian errors while bootstrap intervals do. The results of the Monte Carlo experiments also show that incorporating or not the variability due to parameter estimation makes not difference when building prediction intervals for returns as far as the error distribution is symmetric. However, to construct prediction intervals for future volatilities it is necessary to introduce the uncertainty due to parameter estimation in order to have intervals with coverages close to the nominal values. Although all the results presented in the paper refer to the GARCH(1,1) model, the extension to GARCH(p,q) models is straightforward.

Finally, it is important to mention that the proposed bootstrap prediction intervals for future volatilities are too wide when compared with empirical intervals. Extremely large sample sizes are needed before the bootstrap intervals have the nominal coverages. However, we think that the problem may be in the way the volatility is modelled in GARCH models. Notice that the volatility at time t is observable with information available at time t-1. Therefore, there is not uncertainty associated to one-step ahead volatilities except for the uncertainty associated with parameter estimation. Andersen (2000) shows that volatility diffusion models often used in finance, renders discrete-time strong-form ARCH based models invalid because it is impossible for a discrete return to serve as a sufficient statistic for the innovation to the volatility process. Alternatively, volatility can be modelled by Stochastic Volatility (SV) models as proposed by Harvey et al. (1994) that represent the volatility as an unobservable latent process. It could be worth to investigate the performance of the PRR procedure in the context of SV models.

REFERENCES

[1] ANDERSEN T.G. (2000). Some Reflections on Analysis of High-Frequency Data. *Journal of Business and Economic Statistics* **18**, 146-153.

[2] ANDERSEN, T.G. AND T. BOLLERSLEV (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review* **39**, 885-905.

[3] ANDERSEN, T.G., T. BOLLERSLEV, F.X. DIEBOLD AND P. LABYS (1999). The Distribution of Exchange Rate Volatility. Wharton Financial Institutions Center Working Paper 99-08 and NBER Working Paper 6961.

[4] BAILLIE, R.T. AND T. BOLLERSLEV (1992). Prediction in Dynamic Models with Time Dependent Conditional Variances. *Journal of Econometrics* **52**, 91-113.

[5] BICKEL, P.J. AND D.A. FREEDMAN (1981). Some Asymptotic Theory for the Bootstrap. *Annals of Statistics* **9**, 1196-1217.

[6] BILLINGSLEY, P. (1986). *Probability and Measure*, 2nd ed. New York: Wiley-Interscience.

[7] BOLLERSLEV, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* **31**, 307-327.

[8] BOLLERSLEV, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics* **59**, 542-547.

[9] BOLLERSLEV, T., R.F. ENGLE AND D.B. NELSON (1994). ARCH. The Handbook of Econometrics, Vol. 4, 2959-3038.

[10] BOLLERSLEV, T., AND J.M. WOOLDRIDGE (1992). Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances. *Econometric Reviews* **11(2)**, 143-172.

[11] BOX, G.E.P. AND G.M. JENKINS (1976). *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

[12] CAO, R., M. FEBRERO-BANDE, W. GONZÁLEZ-MANTEIGA, J.M. PRADA-SÁNCHEZ AND I. GARCÍA-JURADO (1997). Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes. *Commun. Statist.* **B 26**(3), 961-978.

[13] DIEBOLD, F.X. (1986). *Empirical Modeling of Exchange Rate Dynamics.* Lectures Notes in Economics and Mathematical Systems. Berlin: Springer-Verlag.

[14] ENGLE, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987-1007.

[15] HARVEY, A.C., E. RUIZ AND N. SHEPHARD (1994). Multivariate Stochastic Variance Models. *Review of Economic Studies* **61**, 247-264.

[16] LUMSDAINE R.L. (1996). Consistency and Asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) Models. *Econometrica* **64**, 575-596.

[17] MIGUEL, J.A. AND P. OLAVE. (1999). Bootstrapping Forecast Intervals in ARCH Models. *Test* **8**, 345-364.

[18] PASCUAL, L., J. ROMO AND E. RUIZ. (1998). Bootstrap Predictive Inference for ARIMA Processes. Working paper 98-86, Universidad Carlos III de Madrid.

[19] PRESS, W.H., B.P. FLANNERY, S.A. TEUKOLSKY AND W.T. VETTERLING (1986). Numerical Recipes, Cambridge University Press, Cambridge.

[20] THOMBS, L.A. AND W.R. SCHUCANY. (1990). Bootstrap Prediction Intervals for Autoregression. *J. Am. Statist. Assoc.* **85**, 486-92.

Table 1. *Monte Carlo results for prediction intervals for returns of* GARCH(1,1) *model with Gaussian innovations*

Lead time	Sample Size	Method	Average Coverage(se)	Coverage below/above	Average Length
1	T	Empirical	95%	2.5%/2.5%	3.82
	300	STD	94.71(.022)	2.64/2.65	3.86(1.00)
		CB	94.45(.024)	2.70/2.85	3.85(1.03)
		PRR	94.52(.023)	2.69/2.79	3.84(.945)
	1000	STD	95.01(.011)	2.50/2.49	3.84(.846)
		CB	94.86(.014)	2.51/2.62	3.83(.858)
		PRR	94.85(.014)	2.53/2.62	3.83(.823)
	3000	STD	95.01(.009)	2.50/2.49	3.85(.908)
		CB	94.89(.011)	2.51/2.60	3.85(.928)
		PRR	94.91(.012)	2.50/2.59	3.84(.910)
10	T	Empirical	95%	2.5%/2.5%	3.90
	300	STD	94.44(.025)	2.78/2.77	3.90(.783)
		CB	94.27(.027)	2.81/2.92	3.91(.833)
		PRR	94.35(.025)	2.78/2.87	3.90(.764)
	1000	STD	94.83(.014)	2.59/2.58	3.90(.588)
		CB	94.80(.016)	2.55/2.65	3.91(.609)
		PRR	94.80(.016)	2.55/2.65	3.91(.576)
	3000	STD	94.86(0.01)	2.58/2.56	3.90(.626)
		CB	94.90(.012)	2.51/2.59	3.92(.646)
		PRR	94.85(.012)	2.53/2.63	3.92(.638)
20	T	Empirical	95%	2.5%/2.5%	3.94
	300	STD	94.30(.026)	2.84/2.86	3.92(.682)
		CB	94.12(.029)	2.87/3.00	3.93(.762)
		PRR	94.23(.024)	2.82/2.95	3.93(.713)
	1000	STD	94.73(.015)	2.62/2.64	3.92(.447)
		CB	94.71(.017)	2.59/2.70	3.94(.475)
		PRR	94.77(.016)	2.55/2.68	3.95(.452)
	3000	STD	94.77(0.01)	2.60/2.63	3.92(.436)
		CB	94.85(.012)	2.50/2.65	3.96(.481)
		PRR	94.83(.012)	2.52/2.65	3.95(.452)

Table 2. *Monte Carlo results for prediction intervals for returns of* GARCH(1,1) *model with Student-5 innovations*

Lead time	Sample Size	Method	Average Coverage(se)	Coverage below/above	Average Length
1	T	Empirical	80%	10%/10%	2.15
	300	STD	83.21(.041)	8.38/8.41	2.42(.869)
		CB	79.58(.046)	10.15/10.27	2.19(.776)
		PRR	79.42(.046)	10.26/10.33	2.16(.672)
	1000	STD	83.74(.027)	8.13/8.14	2.40(.734)
		CB	79.81(.032)	10.05/10.14	2.15(.645)
		PRR	79.77(.031)	10.04/10.19	2.15(.619)
	3000	STD	84.01(.018)	7.98/8.01	2.46(.729)
		CB	79.88(.023)	10.02/10.10	2.20(.655)
		PRR	79.90(.023)	9.98/10.12	2.19(.638)
10	T	Empirical	80%	10%/10%	2.14
	300	STD	84.27(.043)	7.86/7.87	2.51(.716)
		CB	79.58(.045)	10.15/10.26	2.17(.514)
		PRR	79.41(.044)	10.22/10.37	2.14(.432)
	1000	STD	84.63(.031)	7.70/7.67	2.48(.575)
		CB	79.82(.033)	10.06/10.11	2.15(.441)
		PRR	79.70(.032)	10.09/10.21	2.13(.408)
	3000	STD	84.95(.021)	7.53/7.52	2.52(.518)
		CB	79.98(.024)	9.90/10.12	2.17(.428)
		PRR	79.80(.025)	10.01/10.18	2.16(.416)
20	T	Empirical	80%	10%/10%	2.14
	300	STD	84.56(.050)	7.75/7.69	2.55(.669)
		CB	79.52(.048)	10.23/10.24	2.16(.414)
		PRR	79.28(.045)	10.32/10.39	2.14(.370)
	1000	STD	84.97(.034)	7.52/7.50	2.53(.504)
		CB	79.81(.034)	10.07/10.12	2.15(.343)
		PRR	79.71(.032)	10.18/10.11	2.13(.309)
	3000	STD	85.36(.023)	7.34/7.30	2.55(.386)
		CB	79.98(.025)	9.97/10.05	2.16(.287)
		PRR	79.82(.025)	10.02/10.15	2.15(.282)

Table 3. *Monte Carlo results for prediction intervals for returns of* GARCH(1,1) *model with Student-5 innovations*

Lead time	Sample Size	Method	Average Coverage(se)	Coverage below/above	Average Length
1	T	Empirical	99%	0.5%/0.5%	5.92
	300	STD	97.68(.012)	1.16/1.15	4.92(1.76)
		CB	98.42(.012)	.75//.81	6.00(2.52)
		PRR	98.59(.010)	.68/.73	6.07(2.35)
	1000	STD	97.88(.007)	1.07/1.05	4.88(1.49)
		CB	98.78(.007)	.57/.66	5.92(1.94)
		PRR	98.81(.007)	.55/.64	5.95(1.88)
	3000	STD	97.96(.005)	1.03/1.01	4.99(1.48)
		CB	98.86(.005)	.53/.61	6.04(1.93)
		PRR	98.87(.005)	.52/.61	6.04(1.84)
10	T	Empirical	99%	0.5%/0.5%	6.31
	300	STD	97.56(.012)	1.22/1.22	5.12(1.45)
		CB	98.47(.012)	.72/.81	.66(2.39)
		PRR	98.61(.010)	.65/.74	6.48(2.22)
	1000	STD	97.73(.008)	1.14/1.13	5.04(1.17)
		CB	98.77(.007)	.56/.67	6.33(1.73)
		PRR	98.81(.007)	.54/.65	6.39(1.74)
	3000	STD	97.82(.006)	1.10/1.08	5.12(1.05)
		CB	98.90(.005)	.51/.58	6.48(1.56)
		PRR	98.88(.005)	.52/.60	6.43(1.51)
20	T	Empirical	99%	0.5%/0.5%	6.51
	300	STD	97.43(.013)	1.27/1.29	5.19(1.36)
		CB	98.41(.013)	.76/.83	6.52(2.23)
		PRR	98.50(.011)	.71/.79	6.63(2.20)
	1000	STD	97.61(.009)	1.19/1.19	5.13(1.02)
		CB	98.71(.008)	.58/.71	6.56(1.72)
		PRR	98.75(.007)	.57/.68	6.57(1.58)
	3000	STD	97.74(.006)	1.12/1.14	5.18(.784)
		CB	98.85(.005)	.51/.64	6.62(1.31)
		PRR	98.86(.005)	.52/.62	6.59(1.24)

Table 4. *Monte Carlo results for prediction intervals for returns of* GARCH(1,1) *model with Exponential innovations*

Lead time	Sample Size	Method	Average Coverage(se)	Coverage below/above	Average Length
1	T	Empirical	99%	0.5%/0.5%	4.87
	300	STD	96.96(.012)	.00/3.04	4.81(1.85)
		CB	97.63(.041)	1.44/.93	4.88(2.15)
		PRR	99.02(.012)	.13/.85	5.04(2.04)
	1000	STD	97.20(.008)	.00/2.80	4.88(1.79)
		CB	98.37(.023)	.94/.69	4.93(1.97)
		PRR	99.19(.009)	.13/.67	4.98(1.90)
	3000	STD	97.22(.006)	.00/2.78	4.91(1.99)
		CB	98.51(.019)	.85/.64	4.93(1.99)
		PRR	99.20(.008)	.15/.65	4.96(2.20)
10	T	Empirical	99%	0.5%/0.5%	5.70
	300	STD	97.02(.013)	.06/2.92	5.00(1.60)
		CB	97.76(.029)	1.32/.92	5.59(2.36)
		PRR	98.25(.017)	.88/.86	5.74(2.14)
	1000	STD	97.31(.010)	.04/2.65	5.07(1.42)
		CB	98.53(.012)	.79/.69	5.73(1.91)
		PRR	98.64(.010)	.67/.69	5.75(1.88)
	3000	STD	97.35(.006)	.03/2.62	5.09(1.52)
		CB	98.74(.007)	.62/.64	5.72(1.75)
		PRR	98.78(.007)	.58/.63	5.78(2.03)
20	T	Empirical	99%	0.5%/0.5%	5.97
	300	STD	96.96(.014)	.10/2.94	5.10(1.55)
		CB	97.42(.027)	1.58/1.00	5.79(2.44)
		PRR	98.00(.020)	1.07/.93	5.93(2.20)
	1000	STD	97.28(.010)	.08/2.64	5.17(1.26)
		CB	98.35(.013)	.93/.72	5.97(1.83)
		PRR	98.50(.011)	.79/.71	6.03(1.93)
	3000	STD	97.36(.007)	.07/2.57	5.17(1.21)
		CB	98.68(.008)	.67/.65	6.01(1.73)
		PRR	98.75(.007)	.61/.64	6.03(1.80)

Table 5. *Monte Carlo results for prediction intervals of volatilities of* GARCH(1,1) *model with Gaussian innovations*

Lead time	Sample Size	Method	Average Coverage(se)	Coverage below/above	Average Length
1	T	Empirical	95%	2.5%/2.5%	-
	300	CB	-	-	-
		PRR	91.50(.279)	3.40/5.10	.65(.667)
	1000	CB	-	-	-
		PRR	93.70(.243)	3.0/3.30	.32(.249)
	3000	CB	-	-	-
		PRR	94.70(.224)	3.20/2.10	.18(.174)
2	T	Empirical	95%	2.5%/2.5%	.50
	300	CB	57.88(.358)	30.92/11.21	.56(1.01)
		PRR	91.54(.193)	3.63/4.82	.96(1.25)
	1000	CB	70.52(.274)	25.69/3.78	.52(.324)
		PRR	94.19(.122)	2.91/2.90	.68(.433)
	3000	CB	77.46(.223)	19.67/2.87	.51(.321)
		PRR	94.42(.090)	2.91/2.66	.59(.406)
10	T	Empirical	95%	2.5%/2.5%	1.33
	300	CB	75.87(.263)	14.06/10.07	1.33(2.11)
		PRR	87.61(.162)	5.76/6.62	1.56(2.05)
	1000	CB	89.52(.099)	6.56/3.92	1.34(.733)
		PRR	92.57(.074)	3.91/3.52	1.41(.756)
	3000	CB	93.26(.04)	3.80/2.94	1.37(.715)
		PRR	94.17(.03)	2.96/2.87	1.39(.750)
20	T	Empirical	95%	2.5%/2.5%	1.62
	300	CB	75.85(.254)	13.68/10.46	1.58(2.23)
		PRR	85.73(.167)	6.80/7.47	1.79(2.15)
	1000	CB	89.64(.091)	6.18/4.17	1.62(.805)
		PRR	91.83(.074)	4.35/3.81	1.68(.807)
	3000	CB	93.35(.040)	3.27/1.79	1.65(.728)
		PRR	93.90(.035)	3.11/2.99	1.66(.737)

Table 6. *Monte Carlo results for prediction intervals of volatilities of* GARCH(1,1) *model with Student-5 innovations*

Lead time	Sample Size	Method	Average Coverage(se)	Coverage below/above	Average Length
1	T	Empirical	95%	2.5%/2.5%	-
	300	CB	-	-	-
		PRR	89.00(.313)	3.20/7.80	.94(1.30)
	1000	CB	-	-	-
		PRR	93.20(.252)	1.60/5.20	.53(.937)
	3000	CB	-	-	-
		PRR	94.00(.237)	2.20/3.80	.32(.417)
2	T	Empirical	95%	2.5%/2.5%	.58
	300	CB	56.31(.396)	30.80/12.88	.78(2.10)
		PRR	91.87(.201)	2.80/5.33	1.42(2.59)
	1000	CB	65.59(.345)	28.84/5.57	.60(.773)
		PRR	94.83(.123)	1.84/3.32	.92(1.24)
	3000	CB	70.27(.297)	26.59/3.14	.62(.528)
		PRR	95.13(.105)	2.20/2.67	.77(.718)
10	T	Empirical	95%	2.5%/2.5%	1.83
	300	CB	72.57(.294)	16.93/10.50	2.16(4.23)
		PRR	89.22(.155)	4.01/6.77	2.42(3.35)
	1000	CB	85.17(.162)	10.05/4.77	1.85(2.12)
		PRR	92.80(.091)	3.13/4.07	2.02(2.28)
	3000	CB	91.10(.090)	5.80/3.10	1.92(1.43)
		PRR	94.33(.058)	2.71/2.95	1.99(1.43)
20	T	Empirical	95%	2.5%/2.5%	2.25
	300	CB	72.68(.285)	16.28/11.04	2.47(3.54)
		PRR	87.37(.164)	4.95/7.68	2.80(3.55)
	1000	CB	86.36(.142)	8.73/4.91	2.33(2.66)
		PRR	91.93(.091)	3.63/4.44	2.44(2.52)
	3000	CB	91.76(.075)	7.17/2.07	2.36(1.48)
		PRR	94.01(.053)	2.92/3.07	2.42(1.46)

Sample size	Mean	Median	S.D.	Skewness	Kurtosis	Max.	Min.
1065	0.1095	0.1634	1.4199	-0.4661	6.7675*	6.3232	-7.3389
Autocorrelations	r(1)	r(2)	r(3)	r(4)	r(5)	r(10)	r(20)
R_t	0.104*	-0.083*	-0.045	0.011	0.030	0.051	-0.071
$[s.e.]$	[0.045]	[0.044]	[0.045]	[0.045]	[0.042]	[0.046]	[0.041]
R_t^2	0.206	0.187	0.203	0.215	0.164	0.219	0.142

Table 7. Sample moments of returns series of daily IBEX35 index.

Sample size	Mean	Median	S.D.	Skewness	Kurtosis	Max.	Min.
1065	-0.0001	-0.0019	1.3697	-0.3098	5.9578*	5.7889	-7.1571
Autocorrelations	r(1)	r(2)	r(3)	r(4)	r(5)	r(10)	r(20)
y_t	-0.007	-0.066	-0.020	-0.005	0.041	0.066	-0.058
$[s.e.]$	[0.047]	[0.044]	[0.044]	[0.047]	[0.045]	[0.048]	[0.042]
y_t^2	0.273	0.225	0.224	0.274	0.228	0.304	0.179

Table 8. Sample moments of residuals from MA(1) model with interventions.

Sample size	Mean	Median	S.D.	Skewness	Kurtosis	Max.	Min.
1045	0.0067	0.0026	0.9932	-0.2257	3.1470	2.6983	-3.8791
Series/Lag	r(1)	r(2)	r(3)	r(4)	r(5)	r(10)	r(20)
$\widehat{\varepsilon}_t$	-0.007	-0.029	-0.002	0.026	0.013	0.028	-0.032
$[s.e.]$	[0.030]	[0.031]	[0.032]	[0.031]	[0.030]	[0.030]	[0.031]
$\widehat{\varepsilon}_t^2$	-0.007	0.026	0.039	0.026	-0.003	-0.014	0.010

Table 9. Sample moments of standardized residuals.

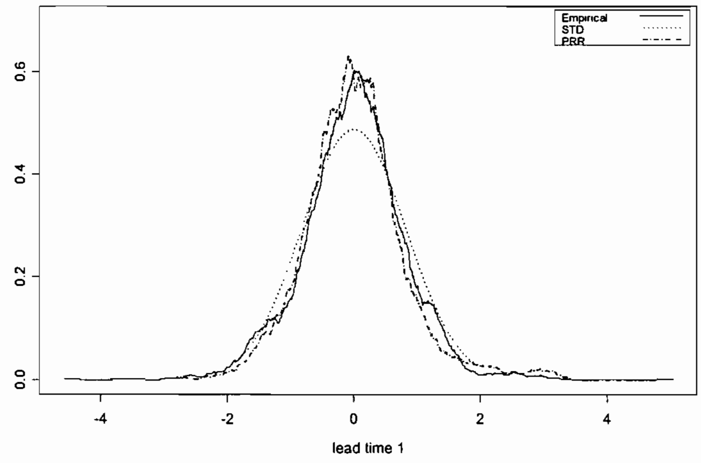


FIG. 1. Estimated kernel densities of one-step ahead predictions of returns of a particular series generated by model (15) with student-5 innovations and T=1000.

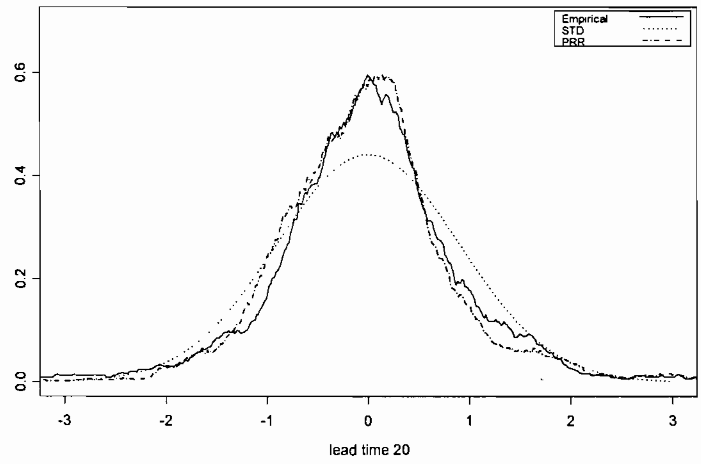


FIG. 2. Estimated kernel densities of twenty-step ahead predictions of returns of a particular series generated by model (15) with student-5 innovations and T=1000.

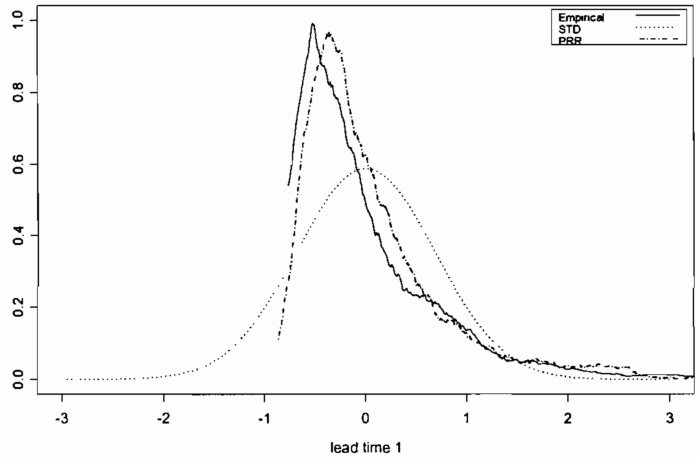


FIG. 3. Estimated kernel densities of one-step ahead predictions of returns of a particular series generated by model (15) with exponential innovations and T=1000.

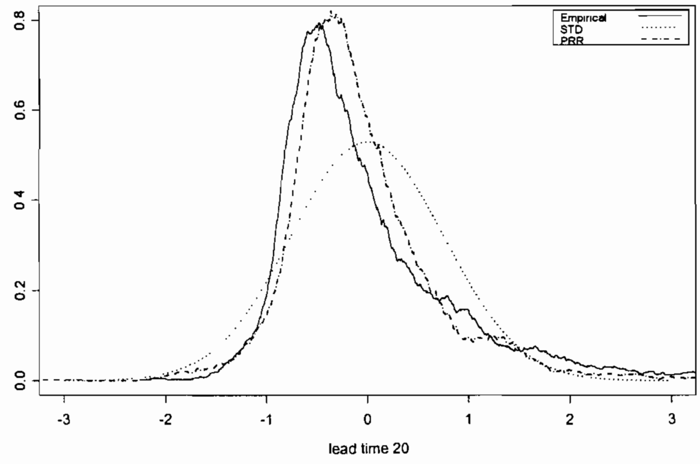


FIG. 4. Estimated kernel densities of twenty-step ahead predictions of returns of a particular series generated by model (15) with exponential innovations and T=1000.

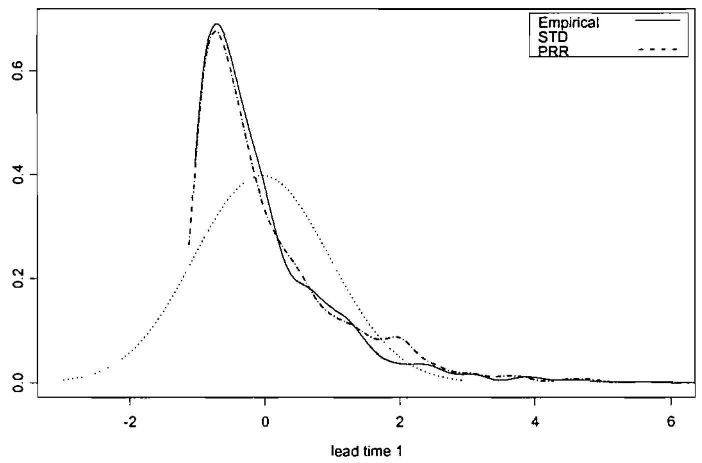


FIG. 5. Estimated kernel densities of one-step ahead predictions of returns of a particular series generated by model (15) with exponential innovations and T=3000.

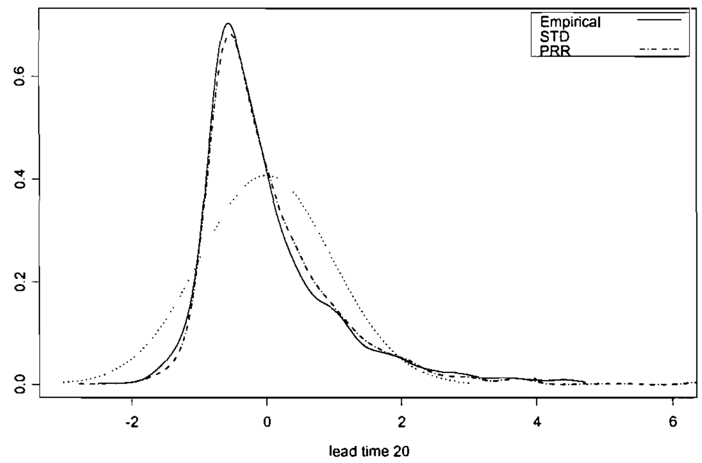


FIG. 6. Estimated kernel densities of twenty-step ahead predictions of returns of a particular series generated by model (15) with exponential innovations and T=3000.

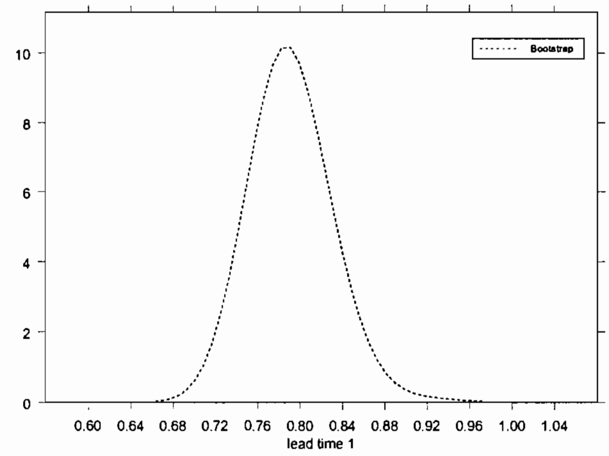


FIG. 7. Estimated kernel densities of one-step ahead predictions of volatilities of a particular series generated by model (15) with Gaussian innovations and T=1000.

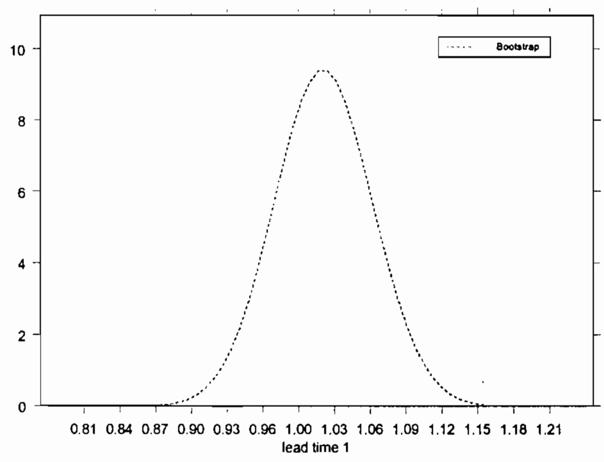


FIG. 8. Estimated kernel densities of one-step ahead predictions of volatilities of a particular series generated by model (15) with Gaussian innovations and T=3000.

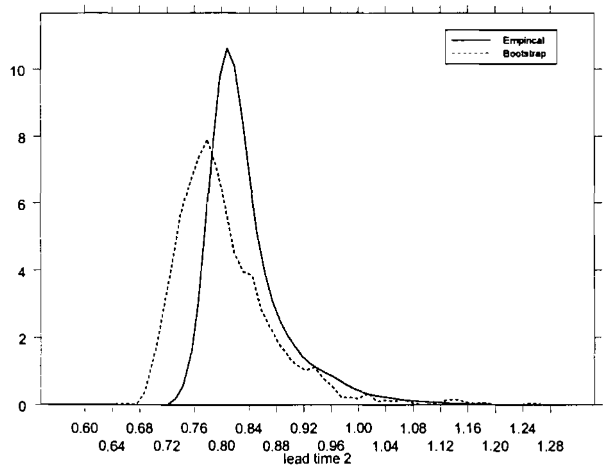


FIG. 9. Estimated kernel densities of two-step ahead predictions of volatilities of a particular series generated by model (15) with Gaussian innovations and T=1000.

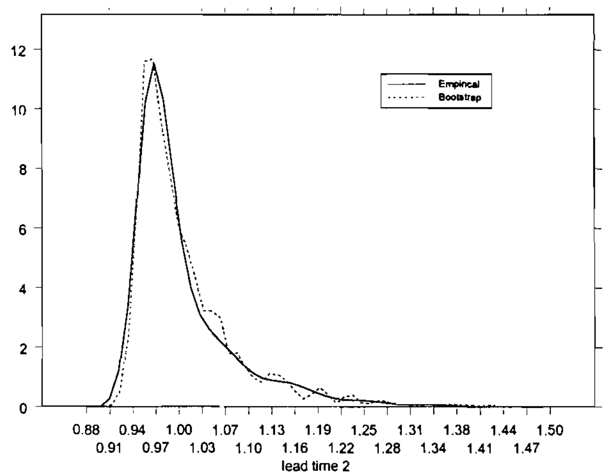


FIG. 10. Estimated kernel densities of two-step ahead predictions of volatilities of a particular series generated by model (15) with Gaussian innovations and T=3000.

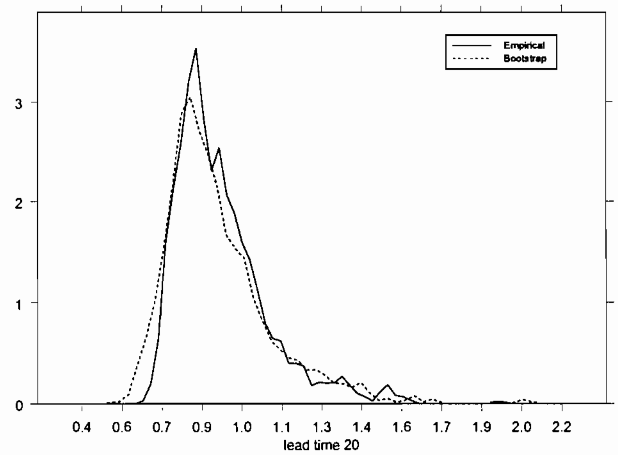


FIG. 11. Estimated kernel densities of twenty-step ahead predictions of volatilities of a particular series generated by model (15) with Gaussian innovations and T=1000.

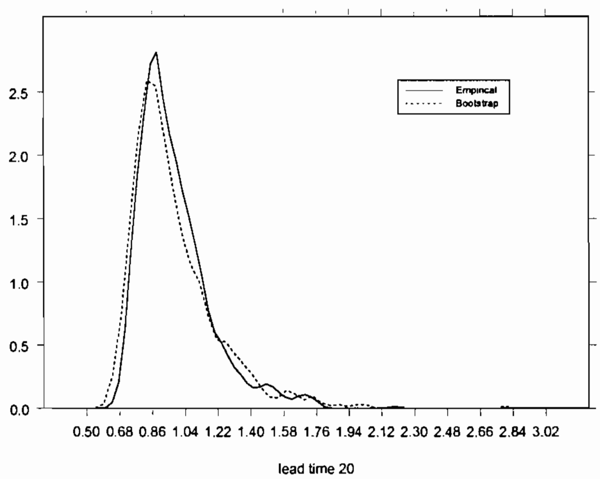


FIG. 12. Estimated kernel densities of twenty-step ahead predictions of volatilities of a particular series generated by model (15) with Gaussian innovations and T=3000.

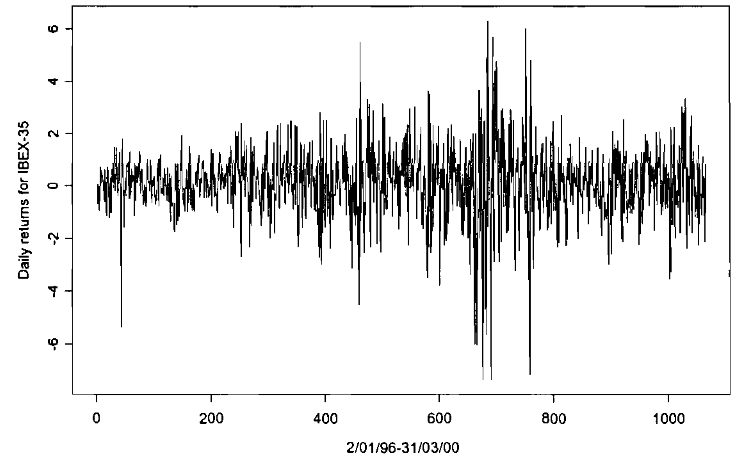


FIG. 13. Returns of IBEX-35 index of Madrid Stock Exchange observed dayly from 2/1/1996 to 31/3/2000

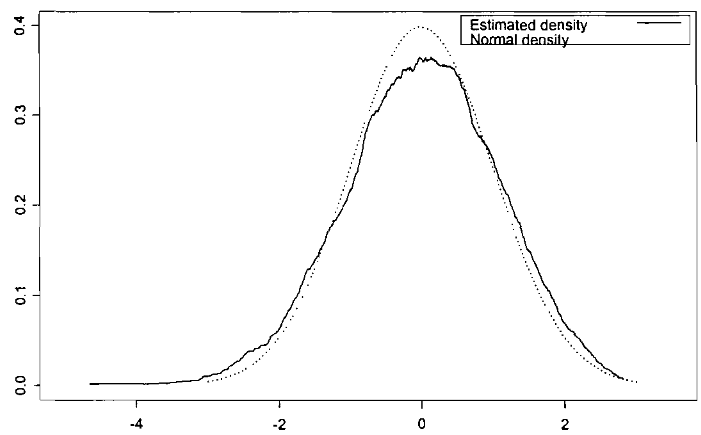


FIG. 14. Estimated kernel density of filtered returns standardized by GARCH(1,1) estimated volatilities and standard normal density

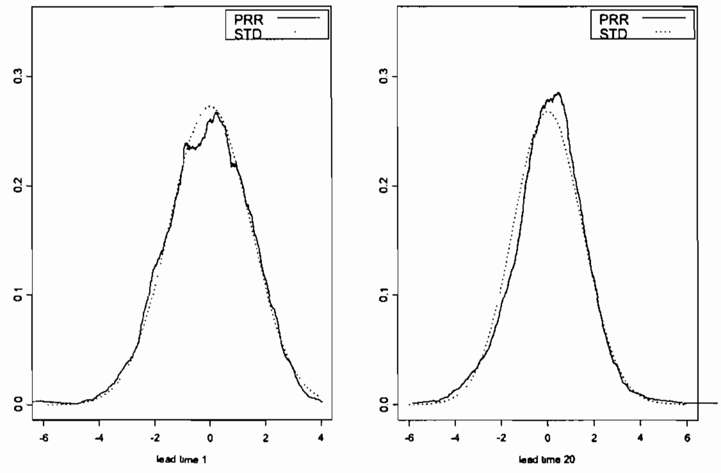


FIG. 15. Estimated kernel densities of one and twenty steps ahead predictions of returns

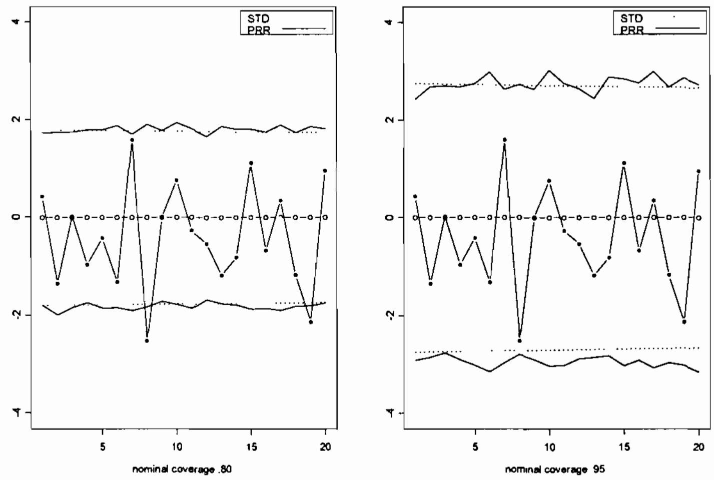


FIG. 16. Prediction intervals of future returns together with real observations (•) and point linear predictions (o)

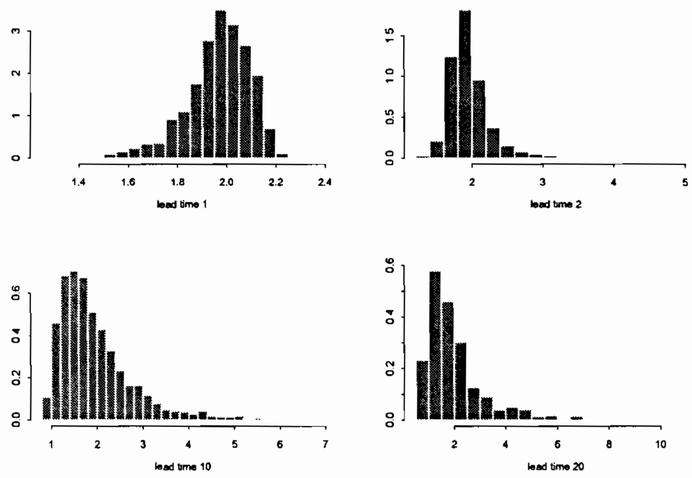


FIG. 17. Histograms of bootstrap predictions of future volatilities of returns

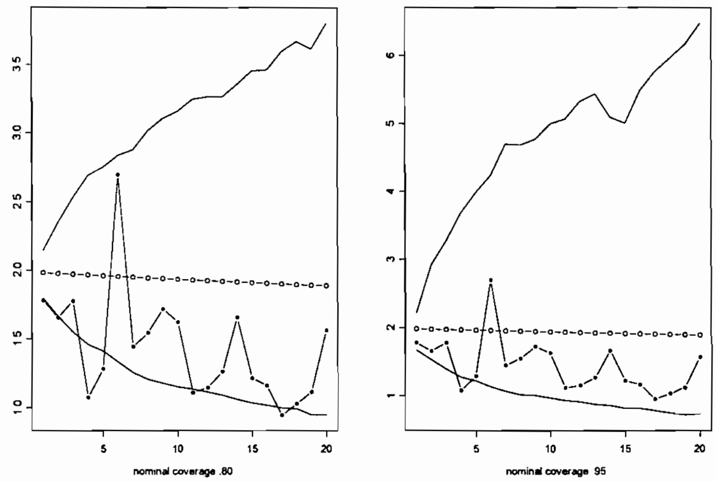


FIG. 18. Bootstrap prediction intervals of future volatilities of returns together with true volatilities (•) and point linear predictions (o)