Journal of Forecasting
J. Forecast. **30**, 51–71 (2011)
Published online 10 November 2010 in Wiley Online Library
(wileyonlinelibrary.com) **DOI**: 10.1002/for.1197

Computationally Efficient Bootstrap Prediction Intervals for Returns and Volatilities in ARCH and GARCH Processes

BEI CHEN,[1]* YULIA R. GEL,[1] N. BALAKRISHNA[2] AND BOVAS ABRAHAM[1]
[1] *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada*
[2] *Department of Statistics, Cochin University of Science and Technology, Cochin, India*

ABSTRACT
We propose a novel, simple, efficient and distribution-free re-sampling technique for developing prediction intervals for returns and volatilities following ARCH/GARCH models. In particular, our key idea is to employ a Box–Jenkins linear representation of an ARCH/GARCH equation and then to adapt a sieve bootstrap procedure to the nonlinear GARCH framework. Our simulation studies indicate that the new re-sampling method provides sharp and well calibrated prediction intervals for both returns and volatilities while reducing computational costs by up to 100 times, compared to other available re-sampling techniques for ARCH/GARCH models. The proposed procedure is illustrated by an application to Yen/U.S. dollar daily exchange rate data.

KEY WORDS financial time series; volatility forecasting; bootstrap; non-Gaussian distribution

INTRODUCTION

Measuring volatility plays an important role in assessing risk and uncertainty in financial markets. One of the core techniques for modeling volatility dynamics in empirical finance is the autoregressive conditional heteroscedastic (ARCH) model, introduced by Engle (1982). The pioneering idea of the ARCH approach is to view volatility as a linear function of previous squared returns. By adding a 'moving average' (MA) part, Bollerslev (1986) proposed incorporating available information on previous volatilities, which resulted in the Generalized ARCH (GARCH) model. There now exists a variety of modifications of the ARCH/GARCH approach: exponential GARCH (Nelson, 1991), nonlinear GARCH (Engle and Ng, 1993), integrated GARCH (Engle and Bollerslev, 1986),

*Correspondence to: Bei Chen, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. E-mail: bchen@math.uwaterloo.ca

fractionally integrated GARCH (Baillie *et al.*, 1996), long memory GARCH (Conrad and Karanasos, 2006), etc., and new extensions continue to appear regularly (see Bera and Higgins, 1993; Shephard, 1996; Tsay, 2002; Taylor, 2005; Bollerslev, 2008; and references therein for an overview). Although the sequence of volatilities is typically unobservable, predicting volatility by ARCH/GARCH models is straightforward due to its functional structure. However, the existing literature mainly focuses on point forecasts of volatility (see Baillie and Bollerslev, 1992; Andersen and Bollerslev, 1998; Andersen *et al.*, 2001; Engle and Patton, 2001; Poon, 2005; and references therein), and relatively little attention has been paid to constructing prediction intervals. Compared to point forecasts, prediction intervals provide extra assessment of the uncertainty associated with the corresponding point forecast, which can better guide risk management decisions. However, construction of prediction intervals requires knowledge of the distribution of the observed data, which is typically unknown in practice. Hence, data are usually assumed to follow some hypothetical distribution, and the resulting prediction interval can be adversely affected by departures from that assumption (Thombs and Schucany, 1990). An alternative is to employ distribution-free re-sampling techniques, e.g., the bootstrap. One of the most popular and efficient bootstrap procedures in a time series context is based on assessing the predictive error distribution by re-sampling residuals from the fitted model (Bühlmann, 2002; Politis, 2003; Härdle *et al.*, 2003). In particular, Miguel and Olave (1999) propose constructing bootstrap-based prediction intervals of returns and volatilities by directly adding re-sampled residuals from an ARCH model to the respective point forecasts. Reeves (2005) suggests including an additional step of re-estimating the ARCH parameters for each bootstrapped realization of returns, which enables uncertainty in sample parameter estimates to be incorporated. Pascual *et al.* (2006) combine and further extend these procedures by developing prediction intervals for both returns and volatilities from GARCH models, and the obtained prediction intervals are found to be well-calibrated, i.e., the number of observed data falling within a prediction interval coincides with the declared coverage. However, the discussed procedures are based on maximum likelihood (ML) estimation of ARCH/GARCH parameters and, hence, are computationally expensive. In this paper, we propose a novel, simple and efficient sieve-based bootstrap procedure to construct prediction intervals of returns and volatilities in ARCH/GARCH processes. The sieve bootstrap is a re-sampling technique that is widely utilized in linear time series modeling due to its efficiency, low computational costs and non-restrictive assumptions (Kreiss, 1988; Bühlmann, 1997; Politis, 2003; Härdle *et al.*, 2003; Pascual *et al.*, 2004). The idea of the sieve bootstrap is to approximate an observed process by a linear model, typically autoregressive (AR), and to generate 'new' realizations from the same model but with the re-sampled innovations. Notice that ARCH/GARCH equations can also be represented as AR/ARMA processes from the Box-Jenkins family of models. In particular, the squared return from an ARCH/GARCH model is a linear process that follows an AR/ARMA equation (Tsay, 2002; Box *et al.*, 2008). Hence, we can also adopt a sieve bootstrap procedure for the ARCH/GARCH case, i.e. develop prediction intervals for squared returns and then apply probability transformations to construct the required prediction intervals for returns and volatility. Since our approach involves only estimation of AR/ARMA models by linear Least Squares (LS), our computational costs are very low, while constructed prediction intervals for both returns and volatility are competitive in terms of coverage and sharpness, compared to other available techniques. In the next section, we discuss the models, assumptions and the new proposed bootstrap algorithm. In section 3, we present an extensive simulation study of various ARCH and GARCH models and distributions for innovation processes. We illustrate our new method by applying it to Yen/U.S. dollar daily exchange rate data in section 4. The paper is concluded by discussion in section 5.

DOI: 10.1002/for

ARCH AND GARCH MODELS AS AR AND ARMA MODELS

We start our discussion from a general class of GARCH(p, q) models and then consider ARCH(1) and GARCH(1, 1) processes as examples.

Models and assumptions

Suppose $\{y_t\}_{t=1}^{T}$ is a GARCH(p, q) process, p, $q \geq 1$. For $t = 1, \ldots, T$:

$$y_t = \sigma_t \varepsilon_t \tag{1}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2}$$

where $\{\varepsilon_t\}_{t=1}^{T}$ is a sequence of independent, identically distributed (i.i.d.) random variables with zero mean, unit variance and $E(\varepsilon_t^4) < \infty$; $\{\sigma_t\}_{t=1}^{T}$ is a stochastic process assumed to be independent of $\{\varepsilon_t\}_{t=1}^{T}$; α_0, α_i and β_j are unknown parameters satisfying $\alpha_0 \geq 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$, for $i = 1, \ldots, p$ and $j = 1, \ldots, q$. Let $m = \max(p, q)$. Throughout this paper, we assume that $\{y_t\}_{t=1}^{T}$ is weakly stationary, i.e. $\Sigma_{i=1}^{m}(\alpha_i + \beta_i) < 1$ is satisfied, where $\alpha_i = 0$ for $i > p$ and $\beta_i = 0$ for $i > q$ (Tsay, 2002). Further, we assume that the strict stationarity conditions of $\{y_t\}_{t=1}^{T}$ given in Bougerol and Picard (1992a,b) hold.[1] Note that in financial contexts $\{y_t\}_{t=1}^{T}$ and $\{\sigma_t\}_{t=1}^{T}$ are referred to as return and volatility processes, respectively. If $q = 0$, $\{y_t\}_{t=1}^{T}$ is an ARCH(p) process. Let us now represent a GARCH(p, q) process in an autoregressive moving average (ARMA) form. If we denote $v_t = y_t^2 - \sigma_t^2$, then

$$y_t^2 = \alpha_0 + \sum_{i=1}^{m} (\alpha_i + \beta_i) y_{t-i}^2 + v_t - \sum_{j=1}^{q} \beta_j v_{t-j} \tag{3}$$

where $\{v_t\}_{t=1}^{T}$ is white noise, but not i.i.d.in general. Under the strict stationary assumption of $\{y_t\}_{t=1}^{T}$, $\{v_t\}_{t=1}^{T}$ is identically distributed. Note that if $q = 0$, formula (3) reduces to an AR(m) model. For example, let us consider the linear forms of two special cases of a GARCH(p, q) process. Suppose $\{y_t\}_{t=1}^{T}$ follows an ARCH(1):

$$y_t = \sigma_t \varepsilon_t \tag{4}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 \tag{5}$$

---

[1] Bougerol and Picard (1992a,b) provide the necessary and sufficient conditions for the existence of a strictly stationary solution of (1) and (2). Let $\tau_n = (\beta_1 + \alpha_1\varepsilon_n^2, \beta_2, \ldots, \beta_{q-1})' \in \mathbb{R}^{q-1}$, $\xi_n = (\varepsilon_n^2, 0, \ldots, 0)' \in \mathbb{R}^{q-1}$, $\alpha = (\alpha_2, \ldots, \alpha_{p-1})' \in \mathbb{R}^{p-2}$. Let

$$A_n = \begin{bmatrix} \tau_n & \beta_q & \alpha & \alpha_p \\ I_{q-1} & 0 & 0 & 0 \\ \xi_n & 0 & 0 & 0 \\ 0 & 0 & I_{p-2} & 0 \end{bmatrix}$$

where I_k is a $k \times k$-identity matrix for $k \in \mathbb{Z}^+$. The top Lyapunov exponent γ_L associated with the sequence $\{A_n, -\infty < n < \infty\}$ is $\gamma_L = \inf_{0 \leq n < \infty} \frac{1}{n+1} E \log\|A_0 A_1 \ldots A_n\|$, assuming that $E(\log\|A_0\|) < \infty$. (Here $\|M\| = \sup\{\|M\mathbf{x}\|_d / \|\mathbf{x}\|_d : \mathbf{x} \in \mathbb{R}^d, \mathbf{x} \neq 0\}$ and $\|\blacksquare\|_d$ is the Euclidean norm in $\mathbb{R}^d$.) Bougerol and Picard (1992a,b) show that if $E(\log\|A_0\|) < \infty$ holds, then (1) and (2) have a unique stationary solution if and only if $\gamma_L < 0$.

DOI: 10.1002/for

Then, in view of (3), equations (4) and (5) can be expressed in the AR(1) form:

$$y_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + v_t \tag{6}$$

Note that $\{y_t\}_{t=1}^T$ is both weakly and strictly stationary if $\alpha_1^2 \leq 1/3$ (Box *et al.*, 2008; Tsay, 2002). Now suppose $\{y_t\}_{t=1}^T$ is a GARCH(1, 1) process given by

$$y_t = \sigma_t \varepsilon_t \tag{7}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{8}$$

Then (7) and (8) can be rewritten as the ARMA(1,1) form:

$$y_t^2 = \alpha_0 + (\alpha_1 + \beta_1) y_{t-1}^2 + v_t - \beta_1 v_{t-1} \tag{9}$$

assuming that $\alpha_1 + \beta_1 < 1$ to ensure the weak stationarity of $\{y_t\}_{t=1}^T$ (Box *et al.*, 2008; Tsay, 2002). Nelson (1990) showed that $\{y_t\}_{t=1}^T$ is also strictly stationary if $E[\log(\beta_1 + \alpha_1 \varepsilon_t)] < 1$. The linear representation of the GARCH(p, q) process allows us to utilize the sieve bootstrap algorithm to construct prediction intervals for returns and volatilities.

Sieve bootstrap procedure of GARCH(p, q) process

The sieve bootstrap is proposed for estimating the distribution of a statistical quantity within a class of linear processes (Kreiss, 1988; Bühlmann, 1997). Given a sample of size T, the idea of the sieve bootstrap is to fit a sequence of AR models of order $p(T)$, where $p(T) \to \infty$ as $T \to \infty$, and then to construct a 'new' bootstrap realization generated from the re-sampled residuals (Grenander, 1981). The asymptotic properties of the sieve bootstrap are studied by Bühlmann (1997), Bickel and Bühlmann (1999), Härdle *et al.* (2003), Politis (2003) and Lahiri (2003). Recently, the sieve bootstrap has been gaining popularity for constructing prediction intervals for linear processes. In particular, Thombs and Schucany (TS) (1990) and Cao *et al.* (1997) consider the performance of sieve bootstrap prediction intervals for finite AR(p) models, while Alonso *et al.* (2002, 2003) extend the sieve bootstrap algorithm to the AR(∞) model with absolutely summable coefficients, and Pascual *et al.* (2004) apply the sieve bootstrap procedure to integrated ARMA (ARIMA) processes. Here we adopt the sieve bootstrap idea for developing prediction intervals for returns and volatility in GARCH(p, q) processes. Let $h = 1, \ldots, s$, $s \geq 1$, be a lead time. Let $\{y_t\}_{t=1}^T$ be an observed sample from (1)–(2), where p and q are assumed to be known.[2] Then, we proceed with the following algorithm to construct prediction intervals for y_{T+h} and σ_{T+h}^2:

- Step 1. Estimate the ARMA coefficients $\hat{\alpha}_0$, $\left(\widehat{\alpha_1 + \beta_1}\right), \ldots, \left(\widehat{\alpha_m + \beta_m}\right)$, $\hat{\beta}_1, \ldots, \hat{\beta}_q$ from the representation (3) using the Least Squares (LS) method. Then, calculate $\hat{\alpha}_i = \left(\widehat{\alpha_i + \beta_i}\right) - \hat{\beta}_i$, for $i = 1, \ldots, p$.
- Step 2. Estimate the residuals $\{\hat{v}_t\}_{t=m+1}^T$ by

$$\hat{v}_t = y_t^2 - \hat{\alpha}_0 - \sum_{i=1}^{m} \left(\widehat{\alpha_i + \beta_i}\right) y_{t-i}^2 + \sum_{j=1}^{q} \hat{\beta}_j \hat{v}_{t-j}, \text{for } t = m+1, \ldots, T \tag{10}$$

Set $\hat{v}_t = 0$, for $t = 1, \ldots, m$.

[2] Note that in practice the order of the GARCH process is unknown. We can select the model order by the Akaike information criterion (AIC) (Akaike, 1974), the Bayesian information criterion (BIC) (Schwarz, 1978) or the corrected AIC (AICc) (Hurvich and Tsai, 1989).

- Step 3. Center the estimated residuals using

$$\tilde{v}_t = \left(\hat{v}_t - \frac{1}{T-m} \sum_{t=m+1}^{T} \hat{v}_t \right), \text{for } t = m+1, \ldots, T \tag{11}$$

The empirical distribution of the centered residuals $\{\tilde{v}_t\}_{t=m+1}^{T}$ is

$$\hat{F}_{v,T}(y) = \sum_{t=m+1}^{T} \mathbf{1}_{\{\tilde{v}_t \le y\}} \tag{12}$$

- Step 4. Sample with replacement from $\hat{F}_{v,T}(y)$ to obtain the bootstrap error process $\{v_t^*\}_{t=1}^{T}$.
- Step 5. Construct a bootstrap sample of squared returns $\{y_t^{2*}\}_{t=1}^{T}$ by

$$y_t^{2*} = \alpha_0 + \sum_{i=1}^{m} \left(\widehat{\alpha_i + \beta_i}\right) y_{t-i}^{2*} + v_t^* - \sum_{j=1}^{q} \beta_j v_{t-j}^* \tag{13}$$

where $y_k^{2*} = \alpha_0 \Big/ \left\{1 - \Sigma_{i=1}^{m} \left(\widehat{\alpha_i + \beta_i}\right)\right\}$ and $v_k^* = 0$ for $k \le 0$. In practice we generate $T + 150$ y_t^{2*} and then discard the first 150 generated values in order to minimize the effect of the initial values. The effect of initial values is negligible asymptotically (Kreiss and Franke, 1992).
- Step 6. Given $\{y_t^{2*}\}_{t=1}^{T}$ from Step 5, estimate the coefficients α_0^*, $\left(\widehat{\alpha_1 + \beta_1}\right)^*, \ldots, \left(\widehat{\alpha_m + \beta_m}\right)^*$, $\hat{\beta}_1^*, \ldots, \hat{\beta}_q^*$ by LS and then calculate $\hat{\alpha}_i^* = \left(\widehat{\alpha_i + \beta_i}\right)^* - \hat{\beta}_i^*$, for $i = 1, \ldots p$. The bootstrap sample of volatility $\{\sigma_t^{2*}\}_{t=1}^{T}$ is obtained by

$$\sigma_t^{2*} = \hat{\alpha}_0^* + \sum_{i=1}^{p} \hat{\alpha}_i^* y_{t-i}^{2*} + \sum_{j=1}^{q} \hat{\beta}_j^* \sigma_{t-j}^{2*} \text{ for } t = m+1, \ldots, T \tag{14}$$

with $\sigma_t^{2*} = \hat{\alpha}_0 / \{1 - \Sigma_{i=1}^{m} (\hat{\alpha}_i + \hat{\beta}_i)\}$, $t = 1, \ldots, m$.
- Step 7. Sample with replacement from $\hat{F}_{v,T}(y)$ to obtain the bootstrap prediction error process $\{v_{T+h}\}_{h=1}^{s}$, where $s \geqslant 1$.
- Step 8. Let $y_{T+h}^* = y_{T+h}$, for $h \leqslant 0$. The h-step-ahead forecast of the squared return is given by

$$y_{T+h}^{2*} = \hat{\alpha}_0^* + \sum_{i=1}^{m} \left(\widehat{\alpha_i + \beta_i}\right)^* y_{T+h-i}^{2*} + v_{T+h}^* - \sum_{j=1}^{q} \hat{\beta}_j^* v_{T+h-j}^* \tag{15}$$

and the h-step-ahead forecast of volatility is obtained by

$$\sigma_{T+h}^{2*} = \hat{\alpha}_0^* + \sum_{i=1}^{p} \hat{\alpha}_i^* y_{T+h-i}^{2*} + \sum_{j=1}^{q} \hat{\beta}_j^* \sigma_{T+h-j}^{2*} \tag{16}$$

for $h = 1, \ldots, s$.
- Step 9. Repeat Steps 4–8 B times, where B is number of bootstrap replicates.

Remark. Note that under the strict stationarity assumption σ_t^2 can be uniquely represented in terms of past observations as

$$\sigma_t^2 = c_0 + \sum_{i=1}^{\infty} c_i y_{t-i}^2 \tag{17}$$

DOI: 10.1002/for

where the c_i's depend on the GARCH parameters α_j and β_j, $j = 1, \ldots, m$ (Berkes *et al.*, 2003). Thus, σ_t^2 is deterministic conditional on the past observations. As addressed in Pascual *et al.* (2006), if the model parameters are known, the one-step ahead volatility is perfectly predictable given $\{y_t\}_{t=1}^T$. The only uncertainty associated with the one-step ahead prediction comes from the parameter estimation whose variability goes to 0 as $T \rightarrow \infty$.

Now, we can use the bootstrap distributions of y_{T+h}^{2*} and σ_{T+h}^{2*} produced by Steps 4–9, i.e. $\hat{F}_{y_{T+h}}^{2*}$ and $\hat{F}_{\sigma_{T+h}}^{2*}$ respectively, to approximate the unknown distributions of y_{T+h}^2 and σ_{T+h}^2, for $h = 1, \ldots, s$. Hence, the $100(1 - \alpha)\%$ prediction interval of y_{T+h}^2 is given by

$$[0, H_{T+h}^*(1-\alpha)], \quad h=1,\ldots,s \tag{18}$$

where $H_{T+h}^*(1 - \alpha)$ is the $(1 - \alpha)$ quantile of $\hat{F}_{y_{T+h}}^{2*}$. The respective $100(1 - \alpha)\%$ prediction interval of y_{T+h} (PI_y^*) is

$$[Q_{T+h}^*(\alpha/2), Q_{T+h}^*(1-\alpha/2)] \tag{19}$$

where $Q_{T+h}^*(\alpha/2) = -\sqrt{H_{T+h}^*(1-\alpha)}$ and $Q_{T+h}^*(1-\alpha/2) = \sqrt{H_{T+h}^*(1-\alpha)}$. Similarly, the $100(1 - \alpha)\%$ bootstrap prediction interval of σ_{T+h}^2 ($PI_{\sigma^2}^*$) is given by

$$[0, K_{T+h}^*(1-\alpha)] \tag{20}$$

where $K_{T+h}^*(1 - \alpha)$ is the $(1 - \alpha)$ quantile of $\hat{F}_{\sigma_{T+h}}^{2*}$.

Remark. Note that to save computing time, similar to the conditional bootstrap of Cao *et al.* (1997) and Miguel and Olave (1999), we can omit the re-estimation Steps 4–6 in our algorithm. We call such a simplified procedure a conditional sieve bootstrap (CSB). However, CSB does not take into account the variance due to parameter estimation. Consequently, the prediction interval of the one-step ahead forecast of volatility does not exist.

NUMERICAL RESULTS

Here we investigate the performance of our method by simulations from various ARCH(p) and GARCH(p, q) models, with $N(0,1)$ and t_5 error distributions. For every combination of model and error distribution, we compare our new unconditional sieve bootstrap (USB) and conditional sieve bootstrap (CSB) with the method proposed by Pascual *et al.* (2006) (PRR) based on the following algorithm:

- Step 1. Simulate the series and generate $R = 1000$ future values y_{T+h} and σ_{T+h}^2, for $h = 1, \ldots, s$. Then, the empirical length of the prediction interval of y_{T+h} is obtained by

$$L_{T+h,y}^e = y_{T+h}^{R(1-\alpha/2)} - y_{T+h}^{R(\alpha/2)} \tag{21}$$

and that of σ_{T+h}^2 is given by

$$L_{T+h,\sigma^2}^e = \sigma_{T+h}^{2,R(1-\alpha/2)} - \sigma_{T+h}^{2,R(\alpha/2)} \tag{22}$$

DOI: 10.1002/for

- Step 2. Compute $B = 1000$ bootstrap forecasts $\{y^{*,b}_{T+h}\}_{b=1}^{B}$ and $\{\sigma^{2*,b}_{T+h}\}_{b=1}^{B}$ and then construct the $100(1-\alpha)\%$ PI^*_y and $PI^*_{\sigma^2}$, $h = 1, \ldots, s$. The lengths of PI^*_y and $PI^*_{\sigma^2}$ are obtained respectively by

$$L^{*,M}_{T+h,y} = Q^*_{T+h}(1-\alpha/2) - Q^*_{T+h}(\alpha/2) \tag{23}$$

and

$$L^{*,M}_{T+h,\sigma^2} = K^*_{T+h}(1-\alpha) \tag{24}$$

- Step 3. Estimate the coverage of PI^*_y and $PI^*_{\sigma^2}$ using

$$C^{*,M}_{T+h,y} = \frac{1}{R}\sum_{r=1}^{R} \mathbf{1}_{\{Q^*_{T+h}(\alpha/2) \le y^{*,r}_{T+h} \le Q^*_{T+h}(1-\alpha/2)\}} \tag{25}$$

and

$$C^{*,M}_{T+h,\sigma^2} = \frac{1}{R}\sum_{r=1}^{R} \mathbf{1}_{\{0 \le \sigma^{2*,r}_{T+h} \le K^*_{T+h}(1-\alpha/2)\}} \tag{26}$$

for $h = 1, \ldots, s$. Note that the coverage is defined as the proportion of future values lying within the prediction interval.
- Step 4. Repeat Steps 1–3 $MC = 1000$ times. Compute the average and the standard error of the coverage of PI^*_y by

$$\bar{C}^*_{T+h,y} = \sum_{M=1}^{MC} C^{*,M}_{T+h,y}/MC \tag{27}$$

$$s.e(\bar{C}^*_{T+h,y}) = \left\{\sum_{M=1}^{MC} (C^{*,M}_{T+h,y} - \bar{C}^*_{T+h,y})^2/MC\right\}^{\frac{1}{2}} \tag{28}$$

and those of the length of PI^*_y by

$$\bar{L}^*_{T+h,y} = \sum_{M=1}^{MC} L^{*,M}_{T+h,y}/MC \tag{29}$$

$$s.e(\bar{L}^*_{T+h,y}) = \left\{\sum_{M=1}^{MC} (L^{*,M}_{T+h,y} - \bar{L}^*_{T+h,y})^2/MC\right\}^{\frac{1}{2}} \tag{30}$$

Similarly, the average and the standard error of the coverage of $PI^*_{\sigma^2}$ are given by

$$\bar{C}^*_{T+h,\sigma^2} = \sum_{M=1}^{MC} C^{*,M}_{T+h,\sigma^2}/MC \tag{31}$$

$$s.e(\bar{C}^*_{T+h,\sigma^2}) = \left\{\sum_{M=1}^{MC} (C^{*,M}_{T+h,\sigma^2} - \bar{C}^*_{T+h,\sigma^2})^2/MC\right\}^{\frac{1}{2}} \tag{32}$$

DOI: 10.1002/for

and those of the length of $PI^*_{\sigma^2}$ are given by

$$\bar{L}^*_{T+h,\sigma^2} = \sum_{M=1}^{MC} L^{*,M}_{T+h,\sigma^2}/MC \tag{33}$$

$$s.e(\bar{L}^*_{T+h,\sigma^2}) = \left\{\sum_{M=1}^{MC}(L^{*,M}_{T+h,\sigma^2} - \bar{L}^*_{T+h,\sigma^2})^2/MC\right\}^{\frac{1}{2}} \tag{34}$$

We compare the performance of USB, CSB and PRR in terms of $\bar{C}^*_{T+h,y}$, $\bar{L}^*_{T+h,y}$, $\bar{C}^*_{T+h,\sigma^2}$, $\bar{L}^*_{T+h,\sigma^2}$ and CPU time, based on the following three models:[3]

- Model 1: ARCH(1)

$$y_t = \sigma_t \varepsilon_t, \tag{35}$$

$$\sigma_t^2 = 0.1 + 0.4 y_{t-1}^2 \tag{36}$$

- Model 2: ARCH(2)

$$y_t = \sigma_t \varepsilon_t, \tag{37}$$

$$\sigma_t^2 = 0.1 + 0.2 y_{t-1}^2 + 0.15 y_{t-2}^2 \tag{38}$$

- Model 3: GARCH(1, 1)

$$y_t = \sigma_t \varepsilon_t, \tag{39}$$

$$\sigma_t^2 = 0.05 + 0.1 y_{t-1}^2 + 0.85 \sigma_{t-1}^2 \tag{40}$$

where $\{\varepsilon_t\}$ follows either N(0, 1) or t_5. In our study, we set the significance level α to 0.05, which corresponds to a 95% prediction interval (PI).

As shown in Tables I–V, USB and CSB provide competitive coverage for PI^*_y, especially for small and moderate sample sizes, while for larger T all three methods perform similarly. In short-term volatility forecasts, USB outperforms PRR for all ARCH models (see Tables I–IV) and both methods yield equivalent results for GARCH models (see Table V). For longer term volatility forecasts, PRR has a slight edge over USB and CSB for small sample sizes. The performance of all three methods tends to be equivalent for larger samples. On comparing USB and CSB, typically USB is somewhat more precise than CSB across all samples, models and distributions. Note that the empirical lengths of prediction intervals of returns and volatilities shown in Tables I–V are obtained using equations (21) and (22).

Now we compare the three methods in terms of CPU time.[4] Figure 1 presents the dynamic of the estimated CPU time for various sample sizes based on $B = 1000$. Note that our results on CPU time are the average of 100 repetitions. As indicated by Figure 1, CSB and USB substantially outperform

[3] We also investigated the algorithm of Miguel and Olave (1999). Their results are found to be equivalent to CSB in terms of coverage and sharpness. However, the required CPU time is substantially higher than that of CSB. Hence we omit the method of Miguel and Olave from further comparison.
[4] The computations were conducted on the Vidal cluster, which has 2 Operon processors with 4 GB RAM on each computing node.

DOI: 10.1002/for

Table I. Prediction intervals for returns of ARCH(1) model with N(0, 1) errors

Lead time	Sample size	Method	Average coverage for return (SE)	Average length for return (SE)	Average coverage for volatility (SE)	Average length for volatility (SE)
1	*T*	Empirical	95%	1.54	95%	—
	300	PRR	94.53 (0.02)	1.53 (0.49)	92.50 (0.26)	0.20 (0.24)
		USB	94.75 (0.04)	1.56 (0.21)	92.00 (0.27)	0.33 (0.11)
		CSB	94.51 (0.05)	1.57 (0.22)	—	—
	1000	PRR	94.83 (0.01)	1.53 (0.43)	93.80 (0.24)	0.18(0.16)
		USB	94.97 (0.04)	1.56 (0.15)	94.20 (0.23)	0.35(0.07)
		CSB	94.89 (0.05)	1.57 (0.23)	—	—
	3000	PRR	94.92 (0.01)	1.54 (0.41)	92.70 (0.26)	0.18 (0.12)
		USB	95.14 (0.04)	1.55 (0.14)	95.00 (0.22)	0.35 (0.05)
		CSB	95.03 (0.04)	1.54 (0.12)	—	—
10	*T*	Empirical	95%	1.61	95%	0.34
	300	PRR	94.61 (0.02)	1.60 (0.14)	93.43 (0.05)	0.36 (0.11)
		USB	94.84 (0.02)	1.62 (0.16)	91.24 (0.06)	0.33 (0.11)
		CSB	94.78 (0.02)	1.61 (0.16)	91.33 (0.06)	0.33 (0.11)
	1000	PRR	94.75 (0.01)	1.60 (0.09)	94.43 (0.02)	0.36 (0.06)
		USB	95.09 (0.01)	1.63 (0.10)	93.55 (0.03)	0.35 (0.07)
		CSB	95.07 (0.01)	1.62 (0.11)	93.60 (0.03)	0.35 (0.08)
	3000	PRR	94.84 (0.01)	1.60 (0.07)	94.79 (0.02)	0.36 (0.04)
		USB	95.14 (0.01)	1.62 (0.08)	94.32 (0.02)	0.35 (0.05)
		CSB	95.19 (0.01)	1.62 (0.08)	94.30 (0.02)	0.35 (0.05)
20	*T*	Empirical	95%	1.62	95%	0.36
	300	PRR	94.54 (0.02)	1.59 (0.14)	93.36 (0.05)	0.36 (0.11)
		USB	94.81 (0.02)	1.62 (0.16)	91.27 (0.06)	0.33 (0.11)
		CSB	94.86 (0.02)	1.62 (0.17)	91.29 (0.06)	0.33 (0.11)
	1000	PRR	94.82 (0.01)	1.60 (0.09)	94.46 (0.02)	0.36 (0.06)
		USB	95.14 (0.01)	1.63 (0.10)	93.62 (0.03)	0.35 (0.07)
		CSB	95.12 (0.01)	1.63 (0.11)	93.56 (0.03)	0.35 (0.08)
	3000	PRR	94.92 (0.01)	1.60 (0.07)	94.79 (0.02)	0.36 (0.04)
		USB	95.12 (0.01)	1.62 (0.08)	94.30 (0.02)	0.35 (0.05)
		CSB	95.11 (0.01)	1.62 (0.07)	94.35 (0.02)	0.35 (0.05)

PRR. In particular, PRR requires 100 times as much CPU time as that of USB for small sample sizes. Remarkably, CSB provides the best performance with only relatively minor loss in terms of sharpness and coverage.

Finally, USB and CSB typically yield some improvements in terms of returns, while PRR generally provides slightly better results for volatilities. With increasing sample size, all three methods perform equivalently. However, USB, and especially CSB, are substantially less computationally demanding. Hence, USB and CSB may be selected as preferred procedures for constructing PIs for returns and volatilities. If only PIs for returns are of interest, then CSB is a better choice.

Remark. In practical applications, a fitted ARCH/GARCH model can exhibit a high degree of persistency. For example, for a case of an GARCH(1, 1) model, high persistency means $\hat{\alpha}_1 + \hat{\beta}_1 \approx 1$,

DOI: 10.1002/for

Table II. Prediction intervals for returns of ARCH(1) model with errors following a *t*-distribution with 5 degrees of freedom, i.e. t_5

Lead time	Sample size	Method	Average coverage for return (SE)	Average length for return (SE)	Average coverage for volatility (SE)	Average length for volatility (SE)
1	*T*	Empirical	95%	1.55	95%	—
	300	PRR	94.56 (0.02)	1.53 (0.54)	85.80 (0.35)	0.22 (0.31)
		USB	94.75 (0.04)	1.55 (0.22)	90.60 (0.29)	0.33 (0.23)
		CSB	94.81 (0.04)	1.55 (0.28)	—	—
	1000	PRR	94.77 (0.01)	1.53 (0.65)	89.10 (0.31)	0.21 (0.65)
		USB	95.00 (0.04)	1.55 (0.19)	93.20 (0.25)	0.33 (0.20)
		CSB	94.90 (0.04)	1.54 (0.17)	—	—
	3000	PRR	94.92 (0.01)	1.54 (0.55)	91.70 (0.28)	0.19 (0.29)
		USB	95.08 (0.03)	1.54 (0.27)	95.00 (0.22)	0.33 (0.09)
		CSB	95.06 (0.03)	1.54 (0.15)	—	—
10	*T*	Empirical	95%	1.62	95%	0.35
	300	PRR	94.59 (0.02)	1.59 (0.21)	92.24 (0.07)	0.37 (0.20)
		USB	95.11 (0.02)	1.66 (0.30)	90.34 (0.08)	0.34 (0.28)
		CSB	95.03 (0.02)	1.65 (0.28)	89.20 (0.10)	0.32 (0.20)
	1000	PRR	94.83 (0.01)	1.60 (0.14)	94.16 (0.03)	0.36 (0.12)
		USB	95.35 (0.01)	1.65 (0.17)	92.58 (0.04)	0.33 (0.20)
		CSB	95.31 (0.01)	1.65 (0.17)	92.26 (0.04)	0.32 (0.12)
	3000	PRR	94.96 (0.01)	1.60 (0.10)	94.75 (0.02)	0.36 (0.07)
		USB	95.41 (0.01)	1.65 (0.12)	93.23 (0.03)	0.33 (0.09)
		CSB	95.44 (0.01)	1.65 (0.12)	93.34 (0.03)	0.33 (0.09)
20	*T*	Empirical	95%	1.64	95%	0.37
	300	PRR	94.56 (0.02)	1.59 (0.21)	92.16 (0.07)	0.37 (0.20)
		USB	95.10 (0.02)	1.66 (0.32)	90.23 (0.08)	0.34 (0.32)
		CSB	95.04 (0.02)	1.65 (0.29)	89.11 (0.10)	0.32 (0.22)
	1000	PRR	94.82 (0.01)	1.60 (0.14)	94.13 (0.03)	0.36 (0.12)
		USB	95.34 (0.01)	1.65 (0.18)	92.52 (0.04)	0.33 (0.20)
		CSB	95.26 (0.01)	1.65 (0.17)	92.16 (0.04)	0.32 (0.11)
	3000	PRR	94.93 (0.01)	1.60 (0.10)	94.65 (0.02)	0.36 (0.07)
		USB	95.42 (0.01)	1.65 (0.12)	93.27 (0.03)	0.33 (0.09)
		CSB	95.43 (0.01)	1.66 (0.12)	93.34 (0.03)	0.33 (0.09)

which may lead to instability of the LS estimation. Although in our studies we have not encountered any stability problems even for a case of $\hat{\alpha}_1 + \hat{\beta}_1$ being 0.95 and 0.981, in practice that might lead to a failure of convergence and inflated standard errors, especially for small and moderate samples. Under such circumstances, we can follow the approach of Kristensen and Linton (2006) and censor the LS estimator at $1 - \varepsilon$ for a small positive ε.

CASE STUDY

In this section we apply the proposed USB and CSB algorithm to construct prediction intervals of returns and volatilities of the daily Yen/U.S. dollar exchange rate, i.e., the number of Yen per U.S.

DOI: 10.1002/for

Table III. Prediction intervals for returns of ARCH(2) model with $N(0, 1)$ errors

Lead time	Sample size	Method	Average coverage for return (SE)	Average length for return (SE)	Average coverage for volatility (SE)	Average length for volatility (SE)
1	T	Empirical	95%	1.52	95%	—
	300	PRR	94.39 (0.02)	1.50 (0.30)	90.30 (0.30)	0.19 (0.10)
		USB	94.59 (0.04)	1.52 (0.16)	91.80 (0.27)	0.28 (0.09)
		CSB	94.65 (0.04)	1.51 (0.15)	—	—
	1000	PRR	94.80 (0.01)	1.52 (0.31)	95.10 (0.22)	0.18 (0.10)
		USB	94.91 (0.03)	1.52 (0.12)	93.60 (0.24)	0.28 (0.05)
		CSB	94.94 (0.03)	1.52 (0.11)	—	—
	3000	PRR	94.85 (0.01)	1.51 (0.32)	93.20 (0.25)	0.17 (0.13)
		USB	94.88 (0.03)	1.52 (0.11)	94.10 (0.24)	0.28 (0.03)
		CSB	94.89 (0.03)	1.52 (0.10)	—	—
10	T	Empirical	95%	1.54	95%	0.25
	300	PRR	94.56 (0.02)	1.54 (0.13)	93.55 (0.06)	0.29 (0.08)
		USB	94.70 (0.02)	1.54 (0.14)	91.92 (0.07)	0.28 (0.09)
		CSB	94.71 (0.02)	1.54 (0.14)	91.15 (0.08)	0.27 (0.09)
	1000	PRR	94.77 (0.01)	1.54 (0.08)	94.30 (0.03)	0.28 (0.04)
		USB	94.86 (0.01)	1.55 (0.08)	93.97 (0.04)	0.28 (0.05)
		CSB	94.87 (0.01)	1.54 (0.08)	93.83 (0.04)	0.28 (0.05)
	3000	PRR	94.84 (0.01)	1.54 (0.07)	94.68 (0.02)	0.27 (0.03)
		USB	94.88 (0.01)	1.54 (0.06)	94.61 (0.02)	0.28 (0.03)
		CSB	94.85 (0.01)	1.54 (0.06)	94.51 (0.02)	0.28 (0.03)
20	T	Empirical	95%	1.55	95%	0.28
	300	PRR	94.59 (0.02)	1.54 (0.12)	93.59 (0.06)	0.29 (0.08)
		USB	94.65 (0.02)	1.54 (0.14)	91.88 (0.07)	0.28 (0.09)
		CSB	94.70 (0.02)	1.54 (0.14)	91.11 (0.08)	0.27 (0.09)
	1000	PRR	94.79 (0.01)	1.54 (0.08)	94.28 (0.03)	0.28 (0.04)
		USB	94.91 (0.01)	1.55 (0.08)	93.94 (0.04)	0.28 (0.05)
		CSB	94.85 (0.01)	1.54 (0.08)	93.88 (0.04)	0.28 (0.05)
	3000	PRR	94.89 (0.01)	1.55 (0.07)	94.66 (0.02)	0.27 (0.03)
		USB	94.90 (0.01)	1.55 (0.06)	94.66 (0.02)	0.28 (0.03)
		CSB	94.87 (0.01)	1.54 (0.06)	94.49 (0.02)	0.28 (0.03)

dollar. In order to avoid modeling particular weekend effects, we exclude all of the observations on Saturdays and Sundays (Andersen *et al.*, 2003). Consequently, our full sample includes the daily average Yen/U.S. exchange rate from March 28th, 1998 to July 28th, 2006: a total of 2175 observations. Figure 2 presents the observed data.

We transform the exchange rate into log returns using

$$y_t = 100 * \log\left(\frac{\text{Yen/U.S. exchange rate in day } t}{\text{Yen/U.S. exchange rate in day } t-1}\right) \tag{41}$$

The new series of returns y_t has a pattern shown in Figure 3, which is stationary and has mean close to zero. Tables VI and VII present summary statistics of y_t.

DOI: 10.1002/for

Table IV. Prediction intervals for returns of ARCH(2) model with errors following a t-distribution with 5 degrees of freedom, i.e. t_5

Lead time	Sample size	Method	Average coverage for return (SE)	Average length for return (SE)	Average coverage for volatility (SE)	Average length for volatility (SE)
1	T	Empirical	95%	1.52	95%	—
	300	PRR	94.56 (0.02)	1.53 (0.55)	87.20 (0.33)	0.23 (0.45)
		USB	94.86 (0.03)	1.55 (0.25)	90.90 (0.29)	0.30 (0.22)
		CSB	94.73 (0.03)	1.53 (0.26)	—	—
	1000	PRR	94.76 (0.01)	1.54 (0.39)	91.20 (0.28)	0.20 (0.25)
		USB	94.89 (0.03)	1.52 (0.15)	92.30 (0.27)	0.28 (0.10)
		CSB	94.77 (0.03)	1.52 (0.15)	—	—
	3000	PRR	94.90 (0.01)	1.52 (0.36)	93.20 (0.25)	0.18 (0.14)
		USB	95.03 (0.03)	1.54 (0.37)	94.50 (0.23)	0.29 (0.07)
		CSB	95.14 (0.03)	1.52 (0.12)	—	—
10	T	Empirical	95%	1.58	95%	0.28
	300	PRR	94.72 (0.02)	1.58 (0.25)	92.40 (0.08)	0.35 (0.52)
		USB	95.13 (0.02)	1.61 (0.25)	91.03 (0.08)	0.30 (0.19)
		CSB	94.91 (0.02)	1.59 (0.34)	88.91 (0.11)	0.30 (0.68)
	1000	PRR	94.84 (0.01)	1.56 (0.15)	94.22 (0.04)	0.32 (0.61)
		USB	95.18 (0.01)	1.59 (0.15)	92.54 (0.05)	0.28 (0.10)
		CSB	95.18 (0.01)	1.59 (0.15)	92.52 (0.05)	0.28 (0.10)
	3000	PRR	94.92 (0.01)	1.56 (0.09)	94.84 (0.02)	0.30 (0.05)
		USB	95.28 (0.01)	1.59 (0.12)	93.93 (0.03)	0.29 (0.10)
		CSB	95.32 (0.01)	1.59 (0.11)	93.78 (0.03)	0.29 (0.07)
20	T	Empirical	95%	1.56	95%	0.30
	300	PRR	94.70 (0.02)	1.57 (0.25)	92.33 (0.08)	0.35 (0.59)
		USB	95.06 (0.02)	1.61 (0.26)	90.96 (0.08)	0.30 (0.20)
		CSB	94.93 (0.02)	1.60 (0.52)	88.91 (0.11)	0.34 (2.05)
	1000	PRR	94.86 (0.01)	1.56 (0.16)	94.18 (0.04)	0.32 (0.50)
		USB	95.22 (0.01)	1.59 (0.15)	92.56 (0.05)	0.28 (0.10)
		CSB	95.22 (0.01)	1.60 (0.16)	92.48 (0.05)	0.28 (0.10)
	3000	PRR	94.97 (0.01)	1.56 (0.10)	94.84 (0.02)	0.30 (0.05)
		USB	95.33 (0.01)	1.60 (0.11)	93.95 (0.03)	0.29 (0.08)
		CSB	95.26 (0.01)	1.59 (0.11)	93.78 (0.03)	0.29 (0.08)

As Table VI indicates, the estimated kurtosis is considerably higher than 3, indicating that the return distribution is leptokurtic. The p-values of the Jarque–Bera test (Jarque and Bera, 1980), the Robust Jarque–Bera test (Gel and Gastwirth, 2008) and the SJ test (Gel *et al.*, 2007) are shown to be less than 0.0001, so there is strong evidence to reject the hypothesis that y_t is Gaussian. Also, autocorrelations of squared returns are highly significant, shown in Table VII. As discussed by West and Cho (1995) as well as Andersen and Bollerslev (1998), a GARCH(1, 1) is a suitable model for y_t.

Next, we partition the full sample into an in-sample estimation set from March 28th, 1998 to June 15th, 2006 and an out-of-sample verification set from June 16th, 2006 to July 28th, 2006. That is,

DOI: 10.1002/for

Table V. Prediction intervals for returns of GARCH(1, 1) model with N(0, 1) errors

Lead time	Sample size	Method	Average coverage for return (SE)	Average length for return (SE)	Average coverage for volatility (SE)	Average length for volatility (SE)
1	T	Empirical	95%	3.81	95%	—
	500	PRR	94.61 (0.02)	3.79 (0.89)	91.50 (0.28)	1.21 (1.37)
		USB	94.76 (0.04)	3.88 (0.45)	91.00 (0.29)	1.38 (0.51)
		CSB	94.69 (0.04)	3.86 (0.46)	—	—
	1000	PRR	94.74 (0.01)	3.82 (0.90)	93.40 (0.25)	1.14 (0.66)
		USB	94.88 (0.03)	3.85 (0.38)	93.40 (0.25)	1.30 (0.44)
		CSB	94.84 (0.04)	3.85 (0.36)	—	—
	3000	PRR	94.87 (0.01)	3.81 (0.86)	94.70 (0.22)	1.07 (0.61)
		USB	94.99 (0.03)	3.86 (0.31)	94.60 (0.23)	1.30 (0.48)
		CSB	94.75 (0.04)	3.88 (0.33)	—	—
10	T	Empirical	95%	3.86	95%	1.66
	500	PRR	94.49 (0.02)	3.88 (0.88)	92.01 (0.08)	1.91 (7.31)
		USB	94.67 (0.03)	3.92 (0.43)	90.12 (0.11)	1.66 (0.64)
		CSB	94.52 (0.03)	3.89 (0.44)	88.29 (0.13)	1.59 (0.60)
	1000	PRR	94.74 (0.02)	3.90 (0.63)	93.36 (0.05)	1.70 (0.81)
		USB	94.78 (0.02)	3.89 (0.35)	92.09 (0.08)	1.65 (0.50)
		CSB	94.69 (0.03)	3.89 (0.35)	91.84 (0.09)	1.64 (0.55)
	3000	PRR	94.86 (0.01)	3.89 (0.61)	94.39 (0.03)	1.67 (0.73)
		USB	94.84 (0.02)	3.89 (0.26)	94.12 (0.05)	1.69 (0.44)
		CSB	94.68 (0.03)	3.92 (0.28)	93.99 (0.05)	1.71 (0.47)
20	T	Empirical	95%	3.92	95%	1.80
	500	PRR	94.32 (0.02)	3.92 (1.50)	91.14 (0.08)	2.44 (20.58)
		USB	94.55 (0.02)	3.92 (0.43)	89.29 (0.10)	1.72 (0.67)
		CSB	94.36 (0.02)	3.90 (0.45)	87.53 (0.12)	1.66 (0.64)
	1000	PRR	94.65 (0.02)	3.92 (0.50)	92.83 (0.05)	1.82 (0.77)
		USB	94.66 (0.02)	3.90 (0.35)	91.31 (0.07)	1.74 (0.55)
		CSB	94.61 (0.02)	3.90 (0.34)	91.26 (0.08)	1.74 (0.62)
	3000	PRR	94.81 (0.01)	3.93 (0.43)	94.24 (0.03)	1.81 (0.60)
		USB	94.73 (0.02)	3.91 (0.25)	93.70 (0.04)	1.80 (0.44)
		CSB	94.66 (0.02)	3.92 (0.24)	93.75 (0.04)	1.81 (0.40)

based on a sample of 2143 observations, we make 31-step ahead predictions. By equation (3), we fit an ARMA(1, 1) model to y_t^2 using LS. The resulting estimated model is given by

$$y_t^2 = 0.006 + 0.9810 y_{t-1}^2 + v_t - 0.9284 v_{t-1} \tag{42}$$

i.e. $\hat{\alpha}_0 = 0.006$, $\hat{\alpha}_1 = 0.0525$ and $\hat{\beta}_1 = 0.9284$. Consistent with the previous literature, the estimate $\hat{\alpha}_1 + \hat{\beta}_1$ is close to unity.

Based on the fitted model (42), we first construct the 95% PIs of returns y_{t+h} from June 16th to July 28th, 2006, using CSB and USB procedures, respectively. Figure 4 presents the estimated distributions of 1-step ahead and 10-step ahead squared returns. Note that since the squared returns are non-negative, each histogram shows a one-sided shape strictly greater than zero. From Figure 4,

DOI: 10.1002/for

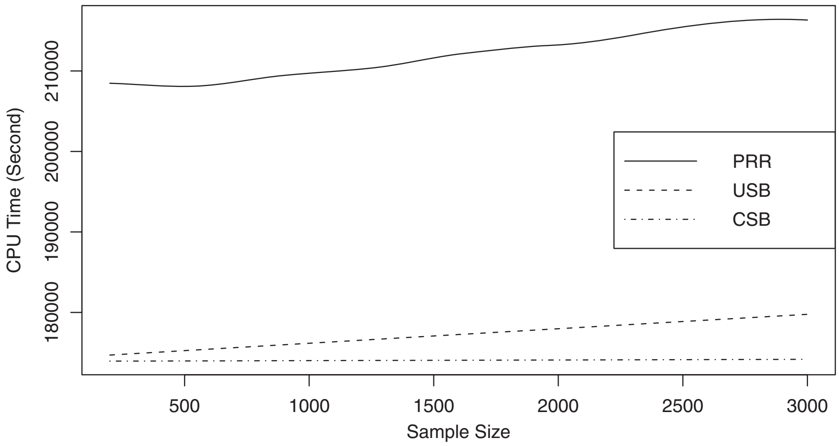

Figure 1. Estimated CPU time for PRR, USB and CSB applied to the GARCH(1,1) process of sample sizes from 200 to 3000

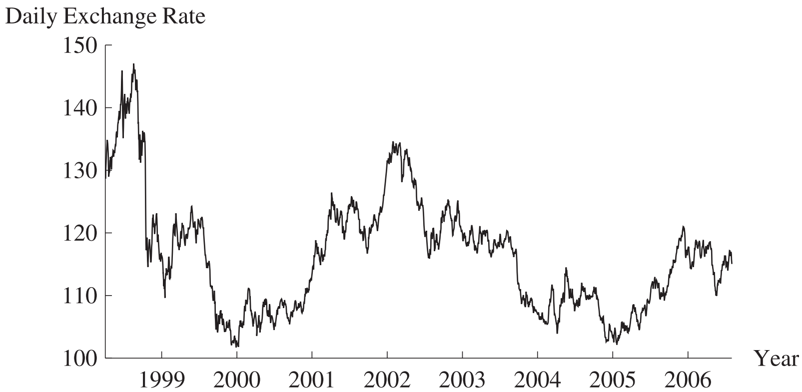

Figure 2. The yen/US daily exchange rates from 28 March 1998 to 28 July 2006

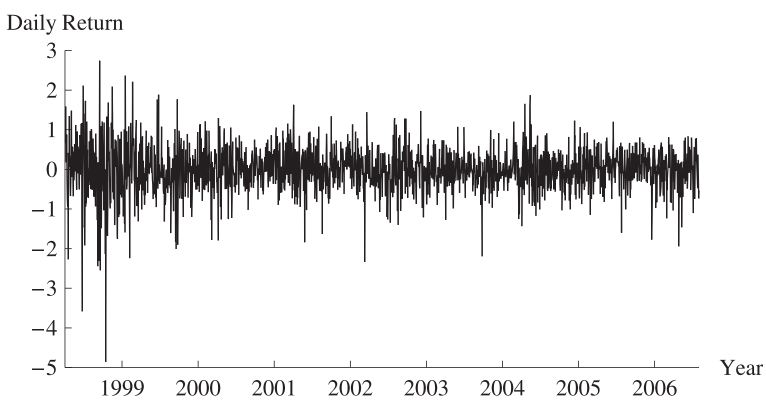

Figure 3. Yen/US daily returns from 28 March 1998 to 28 July 2006

DOI: 10.1002/for

Table VI. Summary statistics for log returns y_t

Mean	Median	SD	Skewness	Kurtosis	Max.	Min.
–0.0024	0.0013	0.5611	–0.6037	8.0743	2.7413	–4.8567

Table VII. Autocorrelations of log returns y_t at different lags

Autocorrelations	$\gamma(1)$	$\gamma(2)$	$\gamma(3)$	$\gamma(10)$	$\gamma(15)$	$\gamma(20)$
y_t	0.2106	0.0086	–0.0573	0.0323	–0.0190	–0.0024
y_t^2	0.0742	0.2204	0.1033	0.0729	0.0916	0.0492

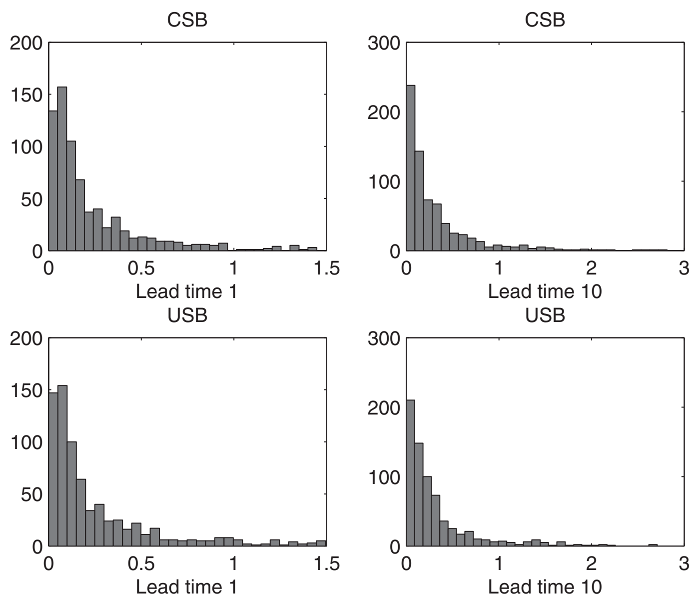

Figure 4. Histograms of bootstrap predictions of the future squared returns

the estimated distributions obtained by CSB are very similar to those by USB. By taking the upper 95% quantiles of the estimated distributions, we acquire the 95% PIs of returns y_{t+h}, $h = 1, \ldots, 31$. Figures 5 and 6 show the 95% PIs of returns, provided by CSB and USB respectively, together with the true values of returns. Notice that the true observations are well covered by the PIs yielded by both CSB and USB.

Finally, we construct the 95% PIs for future volatilities of returns. Figure 7 shows the estimated distributions of future volatilities σ^2_{t+h}, $h = 1, \ldots, 31$. Figure 9 indicates that the distributions of volatilities may be asymmetric. Similarly, we take the upper 95% quantiles of the estimated distributions to construct the PIs of volatilities using CSB and USB. In practice, we do not observe the volatility directly. For verification purposes, we calculate realized volatility from October 9th, 2005 to November 23rd, 2005 based on 5-minute returns using the following equation:

$$\sigma_t^2 = y_{t,1}^2 + \ldots + y_{t,n}^2 \tag{43}$$

DOI: 10.1002/for

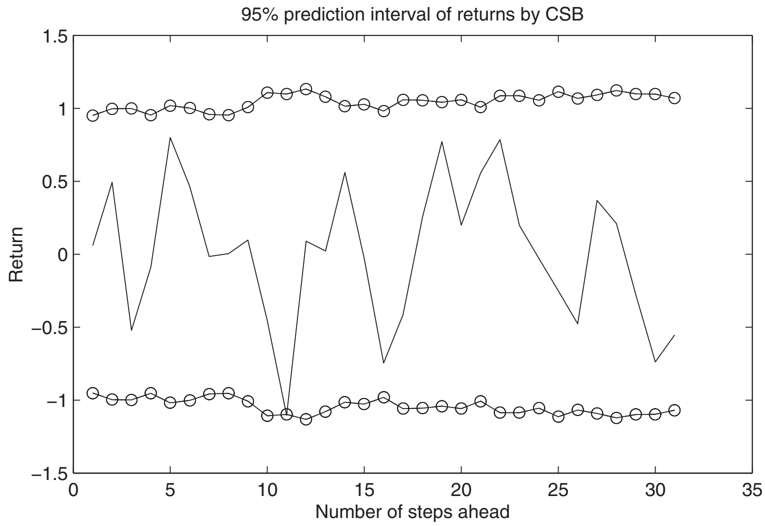

Figure 5. The 95% CSB prediction intervals of returns from 16 June to 18 July 2006

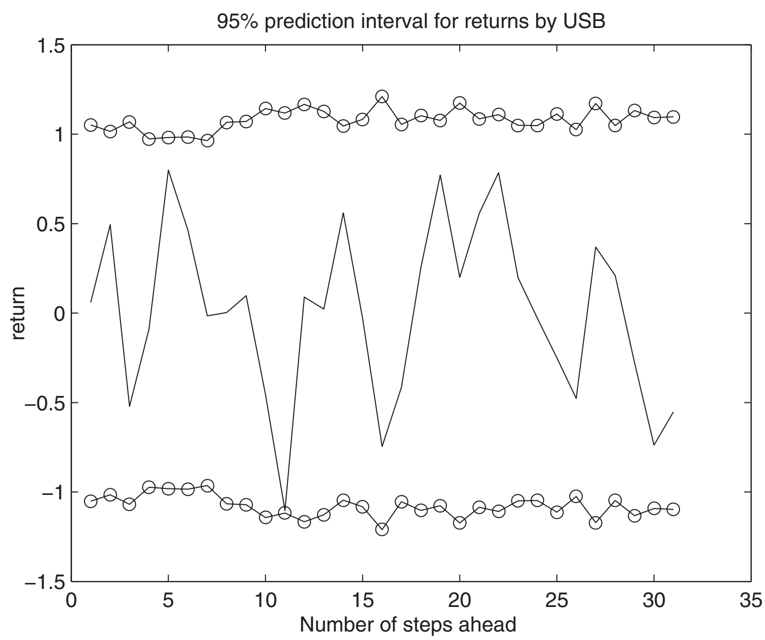

Figure 6. The 95% USB prediction intervals of returns from 16 June to 18 July 2006

where n is the number of observations per day (Andersen and Bollerslev, 1998; Taylor, 2005). Note that n is approximately 268 in our sample. Figures 8 and 9 present the 95% PIs for volatilities together with the realized volatilities. In contrast to the PIs of returns, the 95% PIs of volatilities by USB outperform those of CSB. All of the realized volatilities lie within the 95% USB PIs, but one

DOI: 10.1002/for

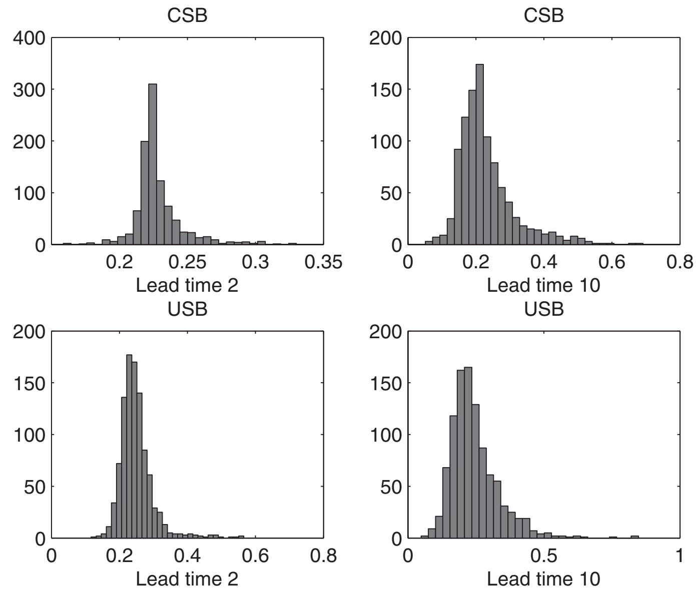

Figure 7. Histograms of bootstrap predictions of the future volatilities of returns

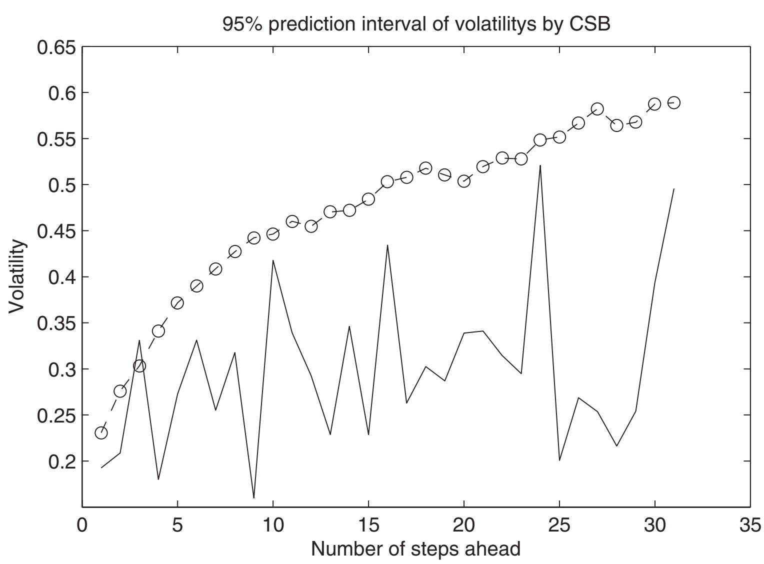

Figure 8. The 95% CSB prediction intervals of volatilities from 16 June to 18 July 2006

DOI: 10.1002/for

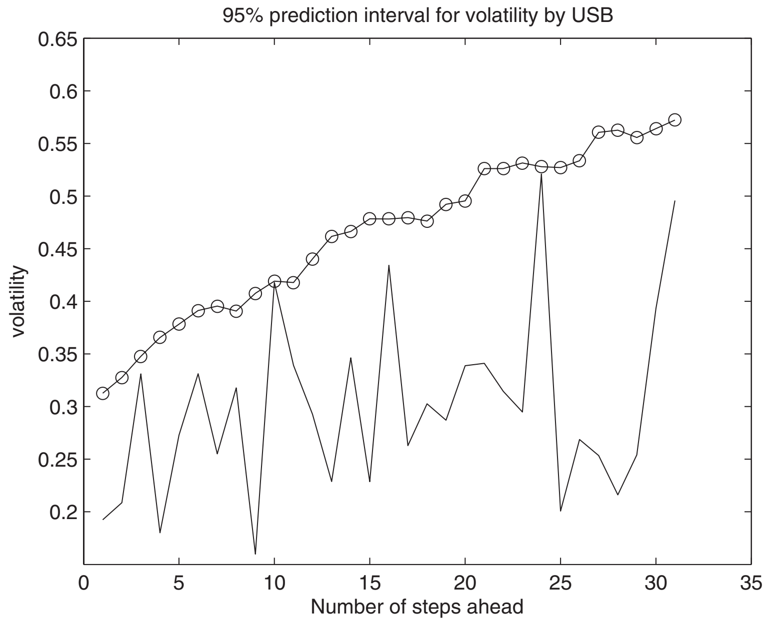

Figure 9. The 95% USB prediction intervals of volatilities from 16 June to 18 July 2006

observation lies outside the boundary in the CSB case. Therefore, the consideration of variance due to parameter estimation is necessary if computational resources is plenty.

DISCUSSION

In this paper we propose a novel, fast and efficient method for constructing prediction intervals for returns and volatilities from ARCH/GARCH models. Our main idea is to transform the non-linear ARCH/GARCH re-sampling problem to a linear framework. In particular, we employ the fact that any stationary ARCH/GARCH model can be represented as a Box-Jenkins model where the squared ARCH/GARCH returns follow a linear AR/ARMA equation. Hence, we can now utilize a sieve bootstrap procedure applied to such an AR/ARMA model. The sieve bootstrap is known to be one of the most efficient and popular re-sampling procedures for linear time series and its asymptotic and computational properties are well investigated. Adapting the sieve bootstrap in an ARCH/GARCH framework allows us to substantially decrease computational costs while providing competitively sharp and well calibrated prediction intervals for both returns and volatilities. The key reason for such an improvement is the fact that a linear AR/ARMA representation of ARCH/GARCH enables us to estimate all the model parameters using recursive Least Squares (LS) or Yule–Walker (YW), which reduces the required computational time by up to a factor of 100, compared to other available re-sampling techniques for ARCH/GARCH models. In addition, such an approach is truly distribution-free and sets minimal restrictions on an ARCH/GARCH innovation process. In the future, we plan to extend the proposed 'linearizing' re-sampling procedure to testing for ARCH/GARCH effects and ARCH/GARCH model selection, as well as to derive asymptotic properties of the sieve bootstrap in an ARCH/GARCH framework.

DOI: 10.1002/for

ACKNOWLEDGEMENTS

We would like to thank Tony Wirjanto for providing the high-frequency data on currency exchange. The research of Yulia R. Gel and Bovas Abraham was supported by grants from the National Science and Engineering Research Council of Canada. Bei Chen was partially supported by the WatRisk scholarship. This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: http://www.sharcnet.ca).

At this occasion of honoring the publication of 'Box–Jenkins', one of the authors of this paper, Bovas Abraham, fondly remembers that he was helping Prof. George Box with checking and proof-reading of the revised edition of this book which came out in 1976.

REFERENCES

Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**(6): 716–723.

Alonso AM, Pẽna D, Romo J. 2002. Forecasting time series with sieve bootstrap. *Journal of Statistical Planning and Inference* **100**: 1–11.

Alonso AM, Pẽna D, Romo J. 2003. On sieve bootstrap prediction intervals. *Statistics and Probability Letters* **65**: 13–20.

Andersen TG, Bollerslev T. 1998. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review* **39**: 885–905.

Andersen TG, Bollerslev T, Diebold FX, Labys P. 2001. The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* **96**: 42–55.

Andersen TG, Bollerslev T, Diebold FX, Labys P. 2003. Modelling and forecasting volatility. *Econometrica* **71**: 579–625.

Baillie RT, Bollerslev T. 1992. Prediction in dynamic models with time-dependent conditional variances. *Journal of Econometrics* **52**: 91–113.

Baillie RT, Bollerslev T, Mikkelsen HO. 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **74**: 3–30.

Bera AK, Higgins ML. 1993. ARCH models: properties, estimation and testing. *Journal of Economic Surveys* **7**: 305–366.

Berkes I, Horváth L, Kokoszka P. 2003. GARCH processes: structure and estimation. *Bernoulli* **9**(2): 201–227.

Bickel P, Bühlmann P. 1999. A new mixing notion and functional central limit theorems for a sieve bootstrap in time series. *Bernoulli* **5**(3): 413–466.

Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**: 307–327.

Bollerslev T. 2008. *Glossary to ARCH (GARCH)*. CREATES Research Paper, 2008–49.

Bougerol P, Picard N. 1992a. Strict stationarity of generalized autoregressive processes. *Annals of Probability* **20**: 1714–1730.

Bougerol P, Picard N. 1992b. Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* **52**: 115–127.

Box GEP, Jenkins GM, Reinsel GC. 2008. *Time Series Analysis: Forecasting and Control* (4th edn). Wiley: Chichester.

Bühlmann P. 1997. Sieve bootstrap for time series. *Bernoulli* **3**: 123–148.

Bühlmann P. 2002. Bootstraps for time series. *Statistical Science* **17**(1): 52–72.

Cao R, Febrero-Bande M, González W, Prada-Sánchez JM, García-Jurado I. 1997. Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes. *Communications in Statistics: Simulation and Computation* **26**: 961–978.

Conrad C, Karanasos M. 2006. The impulse response function of the long memory GARCH process. *Economic Letters* **90**: 34–41.

Engle RF. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**: 987–1007.

DOI: 10.1002/for

Engle RF, Bollerslev T. 1986. Modeling the persistence of conditional variances. *Econometric Reviews* **5**: 1–50.
Engle RF, Ng VK. 1993. Measuring and testing the impact of news on volatility. *Journal of Finance* **48**(5): 1749–1778.
Engle RF, Patton AJ. 2001. What good is a volatility model? *Quantitative Finance* **1**: 237–245.
Gel Y, Gastwirth JL. 2008. The robust Jarque–Bera test of normality. *Economics Letters* **99**(1): 30–32.
Gel Y, Miao W, Gastwirth JL. 2007. Robust directed tests of normality against heavy tailed alternatives. *Computational Statistics and Data Analysis* **51**: 2734–2746.
Grenander U. 1981. *Abstract Inference*. Wiley: New York.
Härdle W, Horowitz JL, Kreiss JP. 2003. Bootstrap methods for time series. *International Statistical Review* **71**: 435–459.
Hurvich CM, Tsai CL. 1989. Regression and time series model selection in small samples. *Biometrika* **76**: 297–307.
Jarque CM, Bera AK. 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters* **6**(3): 255–259.
Kreiss JP. 1988. *Asymptotic statistical inference for a class of stochastic processes*. Habilitationsschrift, Universität Hamburg.
Kreiss JP, Franke J. 1992. Bootstrapping stationary autoregressive moving average models. *Journal of Time Series Analysis* **13**: 297–317.
Kristensen D, Linton O. 2006. A closed-form estimator for the GARCH(1, 1) model. *Econometric Theory* **22**: 323–337.
Lahiri SN. 2003. *Re-sampling Methods for Dependent Data*. Springer: New York.
Miguel JA, Olave P. 1999. Bootstrapping forecast intervals in ARCH models. *Test* **8**(2): 345–364.
Nelson DB. 1990. Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* **6**: 318–344.
Nelson DB. 1991. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**: 347–370.
Pascual L, Romo J, Ruiz E. 2004. Bootstrap predictive inference for ARIMA processes. *Journal of Time Series Analysis* **25**: 449–465.
Pascual L, Romo J, Ruiz E. 2006. Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics and Data Analysis* **50**: 2293–2312.
Politis DN. 2003. The impact of bootstrap methods on time series analysis. *Statistical Science* **28**: 219–230.
Poon SH. 2005. *A Practical Guide to Forecasting Financial Market Volatility*. Wiley: Chichester.
Reeves JJ. 2005. Bootstrap prediction intervals for ARCH models. *Internal Journal of Forecasting* **21**: 237–248.
Schwarz GE. 1978. Estimating the dimension of a model. *Annals of Statistics* **6**(2): 461–464.
Shephard N. 1996. Statistical aspects of ARCH and stochastic volatility models. In *Time Series Models in Econometrics, Finance and Other Fields*, Cox DR, Hinkley DV, Barndorff-Nielsen OE (eds). Chapman & Hall: London; 1–67.
Taylor SJ. 2005. *Asset Price Dynamics and Prediction*. Princeton University Press: Princeton, NJ.
Thombs LA, Schucany WR. 1990. Bootstrap prediction intervals for autoregression. *Journal of the American Statistical Association* **85**: 486–492.
Tsay RS. 2002. *Analysis of Financial Time Series*. Wiley-Interscience: New York.
West KD, Cho D. 1995. The predictive ability of several models of exchange rate volatility. *Journal of Econometrics* **69**: 367–391.

Authors' biographies:

Bei Chen is a Ph.D. student in the Department of Statistics and Actuarial Science at the University of Waterloo, Canada. Her research interests are in time series analysis, nonparametric procedures and econometrics.

Yulia R. Gel is Associate Professor of Statistics in the Department of Statistics and Actuarial Science at the University of Waterloo, Canada. Her research interests are in time series analysis, spatio-temporal modeling and nonparametrics.

N. Balakrishna is Professor of Statistics and Head in the Department of Statistics at the Cochin University of Science and Technology, Cochin, India. His research interests include statistical analysis of time series, inference for stochastic processes and chaotic communications.

DOI: 10.1002/for

Bovas Abraham is the President of the Statistical Society of Canada and is Professor (Emeritus) of Statistics in the Department of Statistics and Actuarial Science at the University of Waterloo, Canada. His research interests are in time series analysis and the application of statistics in business and industry.

Authors' addresses:
Bei Chen, **Yulia R. Gel** and **Bovas Abraham**, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON, Canada N2L 3G1.

N. Balakrishna, Department of Statistics, Cochin University of Science and Technology, Cochin—682 022, India.

DOI: 10.1002/for