Stationary Bootstrap Prediction Intervals for GARCH(p, q)

Eunju Hwang[a], Dong Wan Shin[1,a]

[a]Institute of Mathematical Sciences and Department of Statistics, Ewha Womans University

Abstract

The stationary bootstrap of Politis and Romano (1994) is adopted to develop prediction intervals of returns and volatilities in a generalized autoregressive heteroskedastic (GARCH)(p, q) model. The stationary bootstrap method is applied to generate bootstrap observations of squared returns and residuals, through an ARMA representation of the GARCH model. The stationary bootstrap estimators of unknown parameters are defined and used to calculate the stationary bootstrap samples of volatilities. Estimates of future values of returns and volatilities in the GARCH process and the bootstrap prediction intervals are constructed based on the stationary bootstrap; in addition, asymptotic validities are also shown.

Keywords: GARCH model, stationary bootstrap, prediction, asymptotics.

1. Introduction

For accessing the risks of financial assets, valid predictions of future returns as well as future volatilities are essential. For that purpose, among many time series models, the generalized autoregressive heteroskedastic (GARCH) model is one of the most successful one. Baillie and Bollerslev (1992) studied prediction error distribution in the GARCH(1, 1) model with known parameter values. Andersen and Bollerslev (1998) provided an empirical study to showe that the GARCH model is good for predicting future volatilities. Andersen *et al.* (2001) proposed models and methods to predict realized volatilities. Engle and Patton (2001) compared several volatility models including GARCH(1, 1) for forecasting volatilities.

In order to have more understanding of the risks, we need to know sampling distributions of predictions from which we can construct prediction intervals, Value at Risk (VaR), and others. Even though GARCH models allow the simple constructions of predictions, the mathematical analysis of the sampling distribution would be difficult if the estimated GARCH parameters are plugged-in in constructing the predictions. All the above studies for the prediction assumed given parameter values.

Bootstrapping methods are practical alternatives to mathematical analysis in understanding sampling distributions of predictions that address variations due to estimated parameters. Miguel and Olave (1999) and Reeves (2005) proposed bootstrapping ARCH(p) prediction intervals for future returns. Pascual *et al.* (2006) developed GARCH(1, 1) prediction intervals for returns and volatilities. Chen *et al.* (2011) developed a computationally efficient GARCH(1, 1) bootstrap prediction intervals for returns and volatilities.

However, all the above studies of the bootstrap GARCH predictions did not provide mathematical justification for their bootstrapping methods. Our aim is to construct a mathematically valid bootstrapping method for prediction intervals of future returns and future volatilities. The validity is more

important for risk analysis because an invalid method would provide misleading information about risk. For example an invalid method for VaR would not guarantee the given level (1% or 5%, for example).

We adopt the stationary bootstrapping proposed by Politis and Romano (1994). The stationary bootstrapping is a block resampling method with an increasing random block length that has attracted several authors with recent applications, such as Swensen (2003), Paparoditis and Politis (2005), Parker *et al.* (2006) for unit root tests, and Hwang and Shin (2011, 2012a) for nonparametric analyses. Lahiri (1999), Nordman (2009) and Hwang and Shin (2012b) analyzed new properties of the stationary bootstrap. In our stationary bootstrapping method, the dependence structure of the original data is transferred to the bootstrapping sample by resampling blocks. On the other hand, in the bootstrappings of Miguel and Olave (1999), Reeves (2005), Pascual *et al.* (2006), Chen *et al.* (2011) for GARCH predictions, the dependence structure is transferred to the bootstrapping sample by iterating the GARCH model using normally bootstrapped residuals.

Our method utilizes squares of GARCH processes that have ARMA representations which in turn have long-AR representations. Under the ARMA+AR representation, parameter estimation is linear, which allows us to develop a mathematically valid bootstrapping procedure.

We consider a stationary GARCH(p, q) process defined by (2.1) and (2.2) below. The GARCH model is represented as a form of ARMA. Given a realization of the past observations, the ordinary least square estimators of the ARMA model via long-AR approach are first constructed and residuals are computed. The stationary bootstrap method is applied to generate stationary bootstrap sample of returns, and these data are used to develop the stationary bootstrap estimators, from which the bootstrap sample of volatilities are calculated; in addition, the classical bootstrap residuals are used for estimated *i.i.d.* errors. It is shown that the stationary bootstrap estimators have the same limiting distribution of the ordinary least squares estimators.

Using the stationary bootstrap estimators and the bootstrap samples of returns and volatilities, we obtain bootstrap future values of returns and volatilities of the GARCH process. Slutsky's Theorem and mathematical induction shows that the sampling distribution of the stationary bootstrap future values converges in probability to the distribution of the unknown future value. The stationary bootstrap prediction intervals for returns and volatilities of the GARCH process are constructed by means of bootstrap replicates and quantiles of the Monte Carlo estimates of the bootstrap distributions. By the fundamental weak consistency, the asymptotic validities of the stationary bootstrap prediction intervals are established.

The paper is organized as follows. In Section 2, the stationary bootstrap procedure is described. An algorithm is given for developing the stationary bootstrap estimator of unknown parameters and for constructing the stationary bootstrap prediction intervals of future values of the GARCH process. In Section 3, asymptotic results of the estimators and the prediction intervals are given along with the proofs.

2. Stationary Bootstrapping and Prediction Intervals

We consider a GARCH(p, q) process, $\{y_t\}_{t=1}^{T}$, (for $p, q \geq 1$), satisfying, for $t = 1, \ldots, T$

$$y_t = \sigma_t \epsilon_t, \tag{2.1}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{2.2}$$

where $\{\epsilon_t\}_{t=1}^T$ is a sequence of independent, identically distributed (i.i.d.) random variables with zero mean, unit variance and $E(\epsilon_t^4) < \infty$, and 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 $j = 1, \ldots, q$. Note that the condition $E(\epsilon_t^4) < \infty$ implies $E(y_t^4) < \infty$. In the stochastic process $\{\sigma_t\}_{t=1}^T$ generated by (2.2), σ_t is independent of $\epsilon_t, \epsilon_{t+1}, \ldots, \epsilon_T$. In financial contexts, $\{y_t\}_{t=1}^T$ and $\{\sigma_t\}_{t=1}^T$ are referred to as return and volatility processes. If $q = 0$, $\{y_t\}_{t=1}^T$ is an ARCH(p) process.

In this work we construct the prediction intervals for returns and volatilities of GARCH(p, q) process using the stationary bootstrap. In Section 2.1, the stationary bootstrap procedure of Politis and Romano (1994) is described. In Section 2.2, an algorithm is given to develop the bootstrap prediction intervals. First, the GARCH model is represented in an ARMA form of squared returns and then the coefficients of the ARMA form is estimated by the ordinary squares estimator via a long-AR representation. Using the stationary bootstrap observations, the stationary bootstrap estimators of the ARMA coefficients (and thus the GARCH coefficients) are obtained. From the bootstrap estimated coefficients, the bootstrap sample of volatilities is calculated and then the bootstrap future values of returns and volatilities are obtained. The stationary bootstrap prediction intervals are constructed by the Monte Carlo estimates of the bootstrap distribution of the future values. In Section 3 the consistency results and the large sample validity are given along with the proofs.

2.1. The stationary bootstrap method

Suppose that $\{\mathbf{Y}_t\}$ is a stationary weakly dependent time series taking values in $\mathbb{R}^k$ for some $k \geq 1$. Let $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ be observed. First we define a new time series $\{\mathbf{Y}_{ni} : i \geq 1\}$ by a periodic extension of the observed data set as follows. For each $i \geq 1$, define $\mathbf{Y}_{ni} := \mathbf{Y}_j$ where j is such that $i = qn + j$ for some integer $q \geq 0$. The sequence $\{\mathbf{Y}_{ni} : i \geq 1\}$ is obtained by wrapping the data $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ around a circle, and relabeling them as $\mathbf{Y}_{n1}, \mathbf{Y}_{n2}, \ldots$. Next, for a positive integer l, define the blocks $B(i, l)$, $i \geq 1$ as $B(i, l) = \{\mathbf{Y}_{ni}, \ldots, \mathbf{Y}_{n(i+l-1)}\}$ consisting of l observations starting from $\mathbf{Y}_{ni}$. Bootstrap observations under the stationary bootstrap method are obtained by selecting a random number of blocks from collection $\{B(i, l) : i = 1, \ldots, n, l \geq 1\}$. To do this, we generate random variables $I_1, \ldots, I_n$ and $L_1, \ldots, L_n$ such that conditional on the observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$,

(i) $I_1, \ldots, I_n$ are *i.i.d.* discrete uniform on $\{1, \ldots, n\}$: $P^*(I_1 = i) = 1/n$, $i = 1, \ldots, n$,

(ii) $L_1, \ldots, L_n$ are *i.i.d.* random variables having the geometric distribution with a parameter $\rho \in (0, 1)$: $P^*(L_1 = l) = \rho(1 - \rho)^{l-1}$, $l = 1, 2, \ldots$, where $\rho = \rho(n)$ depends on the sample size n.

(iii) the collections $\{I_1, \ldots, I_n\}$ and $\{L_1, \ldots, L_n\}$ are independent.

Here and in the following, P^* and E^* denote, respectively, the conditional probability and the conditional expectation, given $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$. For notational simplicity, we suppress dependence of the variables $I_1, \ldots, I_n$, $L_1, \ldots, L_n$ and of the parameter ρ on n. We assume that ρ goes to 0 as $n \to \infty$. Under the stationary bootstrap the block length variables $L_1, \ldots, L_n$ are random and the expected block length $E^* L_1$ is ρ^{-1}, which tends to ∞ as $n \to \infty$. Now, a pseudo-time series $\mathbf{Y}_1^*, \ldots, \mathbf{Y}_n^*$ is generated in the following way. Let $\tau = \inf\{k \geq 1 : L_1 + \cdots + L_k \geq n\}$, and select the τ blocks $B(I_1, L_1), \ldots, B(I_\tau, L_\tau)$. Note that there are $L_1 + \cdots + L_\tau$ elements in the resampled blocks $B(I_1, L_1), \ldots, B(I_\tau, L_\tau)$. Arranging these elements in a series and deleting the last $L_1 + \cdots + L_\tau - n$ elements, we get the bootstrap observations $\mathbf{Y}_1^*, \ldots, \mathbf{Y}_n^*$. Conditionally on $\{\mathbf{Y}_1, \ldots, \mathbf{Y}_n\}$, $\{\mathbf{Y}_t^*\}$ is stationary.

2.2. Stationary bootstrap prediction intervals

In order to construct the stationary bootstrap prediction intervals, we give an algorithm with eight steps below. Through the algorithm, we develop the stationary bootstrap estimators of the coefficients of the GARCH model, the bootstrap future values of returns and volatilities, and then the bootstrap prediction intervals.

We assume that the GARCH(p, q) process is strictly stationary. Note that we can find necessary and sufficient condition for the strictly stationarity of the GARCH(p, q) process from Bougerol and Picard (1992a, 1992b).

The GARCH(p, q) process can be represented in an ARMA form as in Chen *et al.* (2011): letting $v_t = y_t^2 - \sigma_t^2$,

$$y_t^2 = \alpha_0 + \sum_{i=1}^{\ell} (\alpha_i + \beta_i) y_{t-i}^2 + v_t - \sum_{j=1}^{q} \beta_j v_{t-j}, \tag{2.3}$$

where $\ell = \max(p, q)$, $\alpha_i = 0$ for $i > p$, $\beta_i = 0$ for $i > q$, and $\{v_t\}_{t=1}^{T}$ is a white noise under the condition $E(y_t^4) < \infty$, but not *i.i.d.* in general. Under the strictly stationary assumption of $\{y_t\}_{t=1}^{T}$, $\{v_t\}_{t=1}^{T}$ is identically distributed.

Here we start from the ARMA model of (2.3) in the following algorithm for developing prediction intervals for returns and volatilities in GARCH(p, q) processes. Let $\{y_{1-\ell}, \ldots, y_0, y_1, \ldots, y_n\}$ be a realization of the GARCH(p, q) process from (2.1) and (2.2).

Step 1. Estimate the ARMA coefficients in (2.3) by $\hat{\alpha}_0, (\widehat{\alpha_1 + \beta_1}), \ldots, (\widehat{\alpha_\ell + \beta_\ell}), \hat{\beta}_1, \ldots, \hat{\beta}_q$, using the least squares estimation of ARMA models similarly to Koreisha and Pukkila (1990) and Kavalieris *et al.* (2003).

We summarize the least squares estimation of ARMA models. Stationary and invertible ARMA(p, q) model can be expressed as an infinite autoregression: thus (2.3) can be written as $y_t^2 = \sum_{j=1}^{\infty} \pi_j y_{t-j}^2 + v_t$. Just as in Koreisha and Pukkila (1990) and Kavalieris *et al.* (2003), first estimate the innovations v_t by $\tilde{v}_t$ through a long-autoregression (LAR) $y_t^2 = \sum_{j=1}^{L} \tilde{\pi}_j y_{t-j}^2 + \tilde{v}_t$ for which π_j values beyond the lag L are assumed to be effectively zero. According to Kavalieris, *et al.* (2003), if

$$\frac{\log n}{2 \log \delta} \le L < H = o\left(\left(\frac{n}{\log n} \right)^{\frac{1}{2}} \right), \tag{2.4}$$

where δ is the modulus of a zero of the generating function $\beta(z) = 1 - \sum_{j=1}^{q} \beta_j z^j$ nearest $|z| = 1$, $\{\eta_t := v_t - \tilde{v}_t\}$ do behave asymptotically like independent random variables; see (2.5) and (2.6) in Kavalieris *et al.* (2003), and the equation in (2.3) results in a regression equation in (2.5) below with moving-average residuals of order q which can be effectively estimated using least squares.

For $t = 1, 2, \ldots, n$

$$y_t^2 - \tilde{v}_t = \alpha_0 + \sum_{i=1}^{\ell} (\alpha_i + \beta_i) y_{t-i}^2 - \sum_{j=1}^{q} \beta_j \tilde{v}_{t-j} + \left(\eta_t - \sum_{j=1}^{q} \beta_j \eta_{t-j} \right), \tag{2.5}$$

where $\{\eta_t\}$ are random white noises. In other words, for $t = 1, 2, \ldots, n$

$$z_t = \mathbf{X}_t \phi + \xi_t,$$

where $z_t = y_t^2 - \tilde{\nu}_t$,

$$\mathbf{X}_t = \left(1, y_{t-1}^2, \ldots, y_{t-\ell}^2, \tilde{\nu}_{t-1}, \ldots, \tilde{\nu}_{t-q}\right),$$
$$\phi = \left(\alpha_0, \alpha_1 + \beta_1, \ldots, \alpha_\ell + \beta_\ell, -\beta_1, \ldots, -\beta_q\right)'$$

and $\xi_t = \eta_t - \sum_{j=1}^q \beta_j \eta_{t-j}$.

Using the defined $z_t, \mathbf{X}_t$ and ξ_t values, we construct the following matrix form:

$$Z = \mathbf{X}\phi + \xi,$$

where $Z = (z_1, \ldots, z_n)'$, $\xi = (\xi_1, \ldots, \xi_n)'$ and

$$\mathbf{X} = \begin{pmatrix} 1 & y_0^2 & \cdots & y_{1-\ell}^2 & \tilde{\nu}_0 & \cdots & \tilde{\nu}_{1-q} \\ 1 & y_1^2 & \cdots & y_{2-\ell}^2 & \tilde{\nu}_1 & \cdots & \tilde{\nu}_{1-q} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & y_{n-1}^2 & \cdots & y_{n-\ell}^2 & \tilde{\nu}_{n-1} & \cdots & \tilde{\nu}_{n-q} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_n \end{pmatrix}.$$

The least squares estimator $\hat{\phi} = (\hat{\alpha}_0, \widehat{(\alpha_1 + \beta_1)}, \ldots, \widehat{(\alpha_\ell + \beta_\ell)}, -\hat{\beta}_1, \ldots, -\hat{\beta}_q)'$ is given by

$$\hat{\phi} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'Z.$$

Calculate $\hat{\alpha}_i = \widehat{(\alpha_i + \beta_i)} - \hat{\beta}_i$, for $i = 1, \ldots, p$.

Step 2. Compute the residuals $\hat{\epsilon}_t$: for $t = 1, \ldots, n$,

$$\hat{\epsilon}_t = \frac{y_t}{\hat{\sigma}_t},$$

where $\hat{\sigma}_t^2 = \hat{\alpha}_0 + \sum_{i=1}^p \hat{\alpha}_i y_{t-i}^2 + \sum_{j=1}^q \hat{\beta}_j \hat{\sigma}_{t-j}^2$ with $\hat{\sigma}_0^2 = \hat{\alpha}_0/(1 - \hat{\alpha}_1 - \hat{\beta}_1)$. Let $\hat{F}_\epsilon$ be the empirical distribution function of the centered and rescaled residuals.

Step 3. Compute the residuals $\hat{\xi}_t$: $\hat{\xi}_t = z_t - \mathbf{X}_t \hat{\phi}$ for $t = 1, \ldots, n$. Let

$$\mathbf{Y}_t = \left(\mathbf{X}_t, \hat{\xi}_t\right) = \left(1, y_{t-1}^2, \ldots, y_{t-\ell}^2, \tilde{\nu}_{t-1}, \ldots, \tilde{\nu}_{t-q}, \hat{\xi}_t\right).$$

Apply the stationary bootstrap method to $\{\mathbf{Y}_1, \ldots, \mathbf{Y}_n\}$, and obtain stationary bootstrap observations $\{\mathbf{Y}_1^*, \ldots, \mathbf{Y}_n^*\}$ with $\mathbf{Y}_t^* = (\mathbf{X}_t^*, \hat{\xi}_t^*) = (1, y_{t-1}^{2*}, \ldots, y_{t-\ell}^{2*}, \tilde{\nu}_{t-1}^*, \ldots, \tilde{\nu}_{t-q}^*, \hat{\xi}_t^*)$.

Step 4. Calculate $\hat{\phi}^* = (\hat{\alpha}_0^*, \widehat{(\alpha_1 + \beta_1)}^*, \ldots, \widehat{(\alpha_\ell + \beta_\ell)}^*, \hat{\beta}_1^*, \ldots, \hat{\beta}_q^*)'$, the stationary bootstrap estimators of the ARMA coefficients with the stationary bootstrap observations in Step 3 as follows:

$$\hat{\phi}^* = (\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{X}^{*'}Z^*,$$

where $\mathbf{X}^* = (\mathbf{X}_1^{*'}, \mathbf{X}_2^{*'}, \ldots, \mathbf{X}_n^{*'})'$ is the $n \times (1 + \ell + q)$ matrix with the stationary bootstrap observations, and $Z^* = \mathbf{X}^*\hat{\phi} + \hat{\xi}^*$ with $\hat{\xi}^* = (\hat{\xi}_1^*, \hat{\xi}_2^*, \ldots, \hat{\xi}_n^*)'$ in Step 3.

Also, calculate $\hat{\alpha}_i^* = \widehat{(\alpha_i + \beta_i)}^* - \hat{\beta}_i^*$ for $i = 1, \ldots, p$.

Step 5. Calculate the bootstrap sample of volatilities $\{\hat{\sigma}_t^{2*} : t = 1, \ldots, n\}$ as follows:

$$\hat{\sigma}_t^{2*} = \hat{\alpha}_0^* + \sum_{i=1}^{p} \hat{\alpha}_i^* y_{t-i}^{2*} + \sum_{j=1}^{q} \hat{\beta}_j^* \hat{\sigma}_{t-j}^{2*}$$

with $\hat{\sigma}_0^{2*} = \hat{\alpha}_0 / (1 - \hat{\alpha}_1 - \hat{\beta}_1)$,

Step 6. Obtain the bootstrap future values of volatilities $\{\hat{\sigma}_{n+k}^{2*} : k = 1, 2, \ldots\}$ and of returns $\{Y_{n+k}^* : k = 1, 2, \ldots\}$ through the following recursions:

$$\hat{\sigma}_{n+k}^{2*} = \hat{\alpha}_0^* + \sum_{i=1}^{p} \hat{\alpha}_i^* Y_{n+k-i}^{2*} + \sum_{j=1}^{q} \hat{\beta}_j^* \hat{\sigma}_{n+k-j}^{2*},$$

$$Y_{n+k}^* = \hat{\sigma}_{n+k}^* \epsilon_{n+k}^*,$$

where $Y_{n+h}^* = y_{n+h}$ for $h \leq 0$ and ϵ_{n+k}^* is a *i.i.d.* random draw from $\hat{F}_\epsilon$ in Step 2.

Step 7. Repeat Steps 3–6 above B times where B is the number of bootstrap replicates, and thus we obtain bootstrap replicates of returns and volatilities, $\{y_{n+k}^{*(1)}, \ldots, y_{n+k}^{*(B)}\}$ and $\{\hat{\sigma}_{n+k}^{2*(1)}, \ldots, \hat{\sigma}_{n+k}^{2*(B)}\}$ for each k.

Step 8. Let $Q_{Y,B}^* = G_{Y,B}^{*-1}$ and $Q_{\sigma^2,B}^* = G_{\sigma^2,B}^{*-1}$ where

$$G_{Y,B}^*(h) = \frac{\#\left(y_{n+k}^{*(b)} \leq h\right)}{B} \quad \text{and} \quad G_{\sigma^2,B}^*(h) = \frac{\#\left(\hat{\sigma}_{n+k}^{2*(b)} \leq h\right)}{B}$$

for $b = 1, \ldots, B$. A $100(1 - \alpha)\%$ bootstrap prediction interval for Y_{n+k} is given by

$$\left[L_{Y,B}^*, \ U_{Y,B}^*\right] = \left[Q_{Y,B}^*\left(\frac{\alpha}{2}\right), \ Q_{Y,B}^*\left(1 - \frac{\alpha}{2}\right)\right]$$

and the $100(1 - \alpha)\%$ bootstrap prediction interval for σ_{n+k}^2 is given by

$$\left[L_{\sigma^2,B}^*, \ U_{\sigma^2,B}^*\right] = \left[Q_{\sigma^2,B}^*\left(\frac{\alpha}{2}\right), \ Q_{\sigma^2,B}^*\left(1 - \frac{\alpha}{2}\right)\right].$$

3. Asymptotic Results

Now we establish the consistency of the stationary bootstrap estimators, and the convergence in distribution of the bootstrap future values as well as the large sample validity of the bootstrap prediction intervals. Proofs are also given.

In the following, for random variables X_n and X, we write $X_n \xrightarrow{p} X$ if X_n converges in probability to X, $X_n \xrightarrow{p^*} X$ if X_n converges to X in conditional probability given the sample $y = \{y_{t-n-\ell}, \ldots, y_t\}$, $X_n \xrightarrow{d} X$ if X_n converges in distribution to X, and $X_n \xrightarrow{d^*} X$ if X_n converges in distribution to X in conditional probability, given the sample $y = \{y_{t-n-\ell}, \ldots, y_t\}$.

Theorem 1. *Let $y = \{y_{t-n-\ell}, \ldots, y_t\}$ be a realization of the GARCH(p, q) process in (2.1) and (2.2). The condition (2.4) in Step 1 is assumed. If $\rho \to 0$ and $n\rho \to \infty$ as $n \to \infty$, then we have, as $n \to \infty$*

(a)

$$\sup_x \left| P^* \left(\sqrt{n} \left[\hat{\phi}^* - \hat{\phi} \right] \leq x \right) - P \left(\sqrt{n} \left[\hat{\phi} - \phi \right] \leq x \right) \right| \xrightarrow{p} 0,$$

and thus

$$\hat{\phi}^* \xrightarrow{p^*} \phi, \quad and \quad \left(\hat{\alpha}_0^*, \hat{\alpha}_1^*, \ldots, \hat{\alpha}_p^*, \hat{\beta}_1^*, \ldots, \hat{\beta}_q^* \right)' \xrightarrow{p^*} \left(\alpha_0, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q \right)', \qquad (3.1)$$

(b) as $n \to \infty$,

$$Y_{t+k}^* \xrightarrow{d^*} Y_{t+k} \quad and \quad \hat{\sigma}_{t+k}^{2*} \xrightarrow{d^*} \sigma_{t+k}^2,$$

(c)

$$\lim_{n \to \infty} \lim_{B \to \infty} P \left[L_{Y,B}^* \leq Y_{t+k} \leq U_{L,B}^* \right] = 1 - \alpha,$$

and

$$\lim_{n \to \infty} \lim_{B \to \infty} P \left[L_{\sigma^2,B}^* \leq \sigma_{t+k}^2 \leq U_{\sigma^2,B}^* \right] = 1 - \alpha.$$

Proof: (a) First, we recall the asymptotic normality of the least squares estimates of ARMA models. The least squares estimator $\hat{\phi}$ satisfies

$$\sqrt{n} \left[\hat{\phi} - \phi \right] = \left(\frac{\mathbf{X}'\mathbf{X}}{n} \right)^{-1} \left(\frac{1}{\sqrt{n}} \right) \mathbf{X}'\xi \xrightarrow{d} N(0, V_\phi),$$

where V_ϕ is the covariance matrix, which is given by $V_\phi = Q^{-1}\Gamma Q^{-1}$ where Q and Γ are such that

$$\left(\frac{\mathbf{X}'\mathbf{X}}{n} \right)^{-1} \xrightarrow{p} Q^{-1}$$

$$\left(\frac{1}{\sqrt{n}} \right) \mathbf{X}'\xi \xrightarrow{d} N(0, \Gamma). \qquad (3.2)$$

Note that Γ involves with the covariance matrix of the moving-average process, which is not a diagonal matrix. In order to show that $\sqrt{n}[\hat{\phi}^* - \hat{\phi}]$ has the same limiting distribution, we observe that

$$\sqrt{n} \left[\hat{\phi}^* - \hat{\phi} \right] = \left(\frac{\mathbf{X}^{*\prime}\mathbf{X}^*}{n} \right)^{-1} \left(\frac{1}{\sqrt{n}} \right) \mathbf{X}^{*\prime}\hat{\xi}^*$$

and we will first show that $\mathbf{X}^{*\prime}\mathbf{X}^*/n$ and $\mathbf{X}'\mathbf{X}/n$ have the same limiting in probability, and secondly show that $\mathbf{X}^{*\prime}\hat{\xi}^*/\sqrt{n}$ and $\mathbf{X}'\xi/\sqrt{n}$ have the same limiting distribution.

We write

$$\mathbf{X}'\mathbf{X} = (\mathbf{X}'_1, \mathbf{X}'_2, \ldots \mathbf{X}'_n) \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_n \end{pmatrix} = \sum_{i=1}^{n} \mathbf{X}_i{}'\mathbf{X}_i$$

and similarly $\mathbf{X}^{*\prime}\mathbf{X}^* = \sum_{i=1}^{n} \mathbf{X}_i^{*\prime}\mathbf{X}_i^*$.

First, we show that $\sum_{i=1}^{n} \mathbf{X}_i^{*\prime}\mathbf{X}_i^*/n$ and $\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i/n$ have the same limiting in probability.

Let $U_{i,r}$ be the sum of $\mathbf{X}'_{nj}\mathbf{X}_{nj}$ in block $B(i, r)$, where $\mathbf{X}_{nj}$ is such that $\mathbf{Y}_{nj} = (\mathbf{X}_{nj}, \hat{\xi}_{nj})$, observations in block $B(i, r) = \{\mathbf{Y}_{nj} : i \le j \le i + r - 1\}$, i.e., $U_{i,r} = \sum_{j=i}^{i+r-1} \mathbf{X}'_{nj}\mathbf{X}_{nj}$. Let $s_\tau = L_1 + L_2 + \cdots + L_\tau$ and use

$$\left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i^{*\prime}\mathbf{X}_i^* - \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i \right| \le \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i^{*\prime}\mathbf{X}_i^* - \frac{1}{n} \sum_{i=1}^{s_\tau} \mathbf{X}_i^{*\prime}\mathbf{X}_i^* \right| + \left| \frac{1}{n} \sum_{i=1}^{s_\tau} \mathbf{X}_i^{*\prime}\mathbf{X}_i^* - \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i \right|. \quad (3.3)$$

We show that the first term of the right-hand side of (3.3) tends to 0 in (conditional) probability;

$$\frac{1}{n} \sum_{i=n+1}^{s_\tau} \mathbf{X}_i^{*\prime}\mathbf{X}_i^* \xrightarrow{p^*} 0.$$

Recalling the definition of τ, and letting $s_{\tau-1} = L_1 + \cdots + L_{\tau-1}$, $R_1 = n - s_{\tau-1}$ and $R = L_\tau - R_1$, we have that $\sum_{i=n+1}^{s_\tau} \mathbf{X}_i^{*\prime}\mathbf{X}_i^*$ is the sum of observations in $B(I_\tau, L_\tau)$, after deleting the first $R_1(= n - s_{\tau-1})$ of them. Note that R, conditional on $(R_1, s_{\tau-1})$, has a geometric distribution with mean $1/\rho$. This follows from the memoryless property of the geometric distribution. Hence, $(1/n) \sum_{i=n+1}^{s_\tau} \mathbf{X}_i^{*\prime}\mathbf{X}_i^*$ is equal in distribution to $(1/n)U_{I,R}$, where I is uniform on $\{1, \ldots, n\}$. It is enough to show that the (conditional) mean and variance of $(1/n)U_{I,R}$ tends to 0.

We have

$$E^*[U_{I,R}|R] = \sum_{i=1}^{n} \frac{1}{n} \left[\sum_{j=i}^{i+R-1} \mathbf{X}'_{nj}\mathbf{X}_{nj} \right] = R \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i \right)$$

and thus

$$\frac{1}{n} E^*[U_{I,R}] = \frac{1}{n} E^* \left[E^*[U_{I,R}|R] \right] = \frac{1}{n\rho} \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i \right) = O_p \left(\frac{1}{n\rho} \right) \xrightarrow{p} 0 \quad (3.4)$$

by (3.2) and since $n\rho \to \infty$. Also we have

$$\frac{1}{n^2} \mathrm{Var}^*[U_{I,R}] = \frac{1}{n^2} E^* \left[\mathrm{Var}^*(U_{I,R}|R) \right] + \frac{1}{n^2} \mathrm{Var}^* \left[E^*(U_{I,R}|R) \right].$$

Its second term $(1/n^2)\mathrm{Var}^*[E^*(U_{I,R}|R)]$ is equal to

$$\frac{1}{n^2} \mathrm{Var}^* \left[R \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i \right) \right] = \frac{1-\rho}{n^2\rho^2} \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i \right) \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\mathbf{X}_i \right)' = O_p \left(\frac{1}{n^2\rho^2} \right) \xrightarrow{p} 0. \quad (3.5)$$

For the first term $(1/n^2)E^*[\text{Var}^*(U_{I,R}|R)]$, consider $\text{Var}^*(U_{I,R}|R) = E^*[U_{I,R}U'_{I,R}|R] =$

$$\sum_{i=1}^{n} \frac{1}{n} \left(\sum_{j=i}^{i+R-1} \mathbf{X}'_{nj}\mathbf{X}_{nj} \right) \left(\sum_{j=i}^{i+R-1} \mathbf{X}'_{nj}\mathbf{X}_{nj} \right)' = \frac{R}{n} \left(\sum_{i=1}^{n} \mathbf{X}'_i\mathbf{X}_i \right) \left(\sum_{i=1}^{n} \mathbf{X}'_i\mathbf{X}_i \right)'$$

and then

$$\frac{1}{n^2} E^* \left[\text{Var}^*(U_{I,R}|R) \right] = \frac{1}{n^3\rho} \left(\sum_{i=1}^{n} \mathbf{X}'_i\mathbf{X}_i \right) \left(\sum_{i=1}^{n} \mathbf{X}'_i\mathbf{X}_i \right)' = O_p \left(\frac{1}{n\rho} \right) \xrightarrow{p} 0. \tag{3.6}$$

Thus the (conditional) variance of $(1/n)U_{I,R}$ tends to 0, and the first term of the right-hand side of (3.3) tends to 0 in (conditional) probability.

Now we show that the second term of the right-hand side of (3.3) tends to 0 in (conditional) probability;

$$\left| \frac{1}{n} \sum_{i=1}^{s_\tau} \mathbf{X}_i^{*\prime}\mathbf{X}_i^* - \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}'_i\mathbf{X}_i \right| \xrightarrow{p^*} 0.$$

Recalling the definition of $U_{i,r} = \sum_{j=i}^{i+r-1} \mathbf{X}'_{nj}\mathbf{X}_{nj}$, we have

$$\frac{1}{n} \sum_{i=1}^{s_\tau} \mathbf{X}_i^{*\prime}\mathbf{X}_i^* = \frac{1}{n} \sum_{i=1}^{\tau} U_{I_i,L_i}.$$

Just as in Politis and Romano (1994) and Hwang and Shin (2012b), since $\tau = n\rho + O_p(\sqrt{n\rho})$, we consider a sequence $m = m_n$ with $m/(n\rho) \to 1$, and it suffices to show that

$$\left| \frac{1}{n} \sum_{i=1}^{m} U_{I_i,L_i} - \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}'_i\mathbf{X}_i \right| \xrightarrow{p^*} 0.$$

The left-hand side is less than or equal to

$$\left| \frac{1}{n} \sum_{i=1}^{m} U_{I_i,L_i} - \frac{\rho}{m} \sum_{i=1}^{m} U_{I_i,L_i} \right| + \left| \frac{\rho}{m} \sum_{i=1}^{m} U_{I_i,L_i} - \rho E^*[U_{I_1,L_1}] \right| + \left| \rho E^*[U_{I_1,L_1}] - \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}'_i\mathbf{X}_i \right|. \tag{3.7}$$

The first term of (3.7) is less than or equal to

$$\left| \sum_{i=1}^{m} U_{I_i,L_i} \right| \left| \frac{1}{n} - \frac{\rho}{m} \right| = \left| \frac{1}{nm} \sum_{i=1}^{m} U_{I_i,L_i} \right| |m - n\rho| = o_p \left(\frac{1}{n\rho} \right) \xrightarrow{p^*} 0,$$

where the last equality holds by (3.4). The second term of (3.7),

$$\rho \left| \frac{1}{m} \sum_{i=1}^{m} U_{I_i,L_i} - E^*[U_{I_1,L_1}] \right| = o_p(\rho) \xrightarrow{p^*} 0,$$

where the last equality holds by the weak law of large numbers of *i.i.d.* sequence $\{U_{I_i,L_i} : i = 1, 2, \dots\}$ as $m \to \infty$. For the third term of (3.7), we calculate $E^*[U_{I_1,L_1}]$. Similarly to above, we have

$$E^*[U_{I_1,L_1}] = E^* \left[E^*(U_{I_1,L_1}|L_1) \right] = E^* \left[L_1 \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}'_i\mathbf{X}_i \right) \right] = \frac{1}{\rho} \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}'_i\mathbf{X}_i \right).$$

Thus the third term of (3.7) is zero. Therefore, $\sum_{i=1}^{n} \mathbf{X}_i^{*\prime} \mathbf{X}_i^* / n$ and $\sum_{i=1}^{n} \mathbf{X}_i^{\prime} \mathbf{X}_i / n$ have the same limiting in probability.

Secondly, we show that $(1/\sqrt{n}) \mathbf{X}^{*\prime} \hat{\xi}^* \xrightarrow{d^*} N(0, \Gamma)$. We write

$$\mathbf{X}^{*\prime} \hat{\xi}^* = (\mathbf{X}_1^{*\prime}, \mathbf{X}_2^{*\prime}, \dots \mathbf{X}_n^{*\prime})(\hat{\xi}_1^*, \hat{\xi}_2^*, \cdots, \hat{\xi}_n^*)' = \sum_{i=1}^{n} \mathbf{X}_i^{*\prime} \hat{\xi}_i^*, \quad \mathbf{X}'\xi = \sum_{i=1}^{n} \mathbf{X}_i'\xi_i,$$

and

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_i^{*\prime} \hat{\xi}_i^* = \frac{1}{\sqrt{n}} \sum_{i=1}^{s_\tau} \mathbf{X}_i^{*\prime} \hat{\xi}_i^* - \frac{1}{\sqrt{n}} \sum_{i=n+1}^{s_\tau} \mathbf{X}_i^{*\prime} \hat{\xi}_i^*.$$

Let $V_{i,r} = \sum_{j=i}^{i+r-1} \mathbf{X}_{nj}' \hat{\xi}_{nj}$, which is related to the observations in block $B(i, r)$. By the same arguments as above, $(1/\sqrt{n}) \sum_{i=n+1}^{s_\tau} \mathbf{X}_i^{*\prime} \hat{\xi}_i^*$ is equal in distribution to $(1/\sqrt{n}) V_{I,R}$, and, instead of (3.4), (3.5) and (3.6), respectively, we have

$$\frac{1}{\sqrt{n}} E^*[V_{I,R}] = \frac{1}{\sqrt{n}} E^*[E^*[V_{I,R}|R]] = \frac{1}{n\rho} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i \right) = O_p\left(\frac{1}{n\rho}\right) \xrightarrow{p} 0$$

$$\frac{1}{n} \mathrm{Var}^*[E^*(V_{I,R}|R)] = \frac{1}{n} \mathrm{Var}^* \left[R\left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i \right) \right] = \frac{1-\rho}{n^2\rho^2} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i \right) \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i \right)' \qquad (3.8)$$

which is equal to $O_p(1/(n^2\rho^2)) \xrightarrow{p} 0$, and

$$\frac{1}{n} E^*[\mathrm{Var}^*(V_{I,R}|R)] = \frac{1}{n\rho} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i \right) \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i \right)' = O_p\left(\frac{1}{n\rho}\right) \xrightarrow{p} 0,$$

and thus we have

$$\frac{1}{\sqrt{n}} \sum_{i=n+1}^{s_\tau} \mathbf{X}_i^{*\prime} \hat{\xi}_i^* \xrightarrow{p^*} 0.$$

Recalling the definition of $V_{i,r} = \sum_{j=i}^{i+r-1} \mathbf{X}_{nj}' \hat{\xi}_{nj}$, we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{s_\tau} \mathbf{X}_i^{*\prime} \hat{\xi}_i^* = \frac{1}{\sqrt{n}} \sum_{i=1}^{\tau} V_{I_i, L_i}.$$

Similarly to above, for a sequence $m = m_n$ with $m/(n\rho) \to 1$, it suffices to show that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{m} V_{I_i, L_i} \xrightarrow{d^*} N(0, \Gamma).$$

The left-hand side is equal to

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{m} V_{I_i, L_i} - \sqrt{\frac{\rho}{m}} \sum_{i=1}^{m} V_{I_i, L_i} \right) + \left(\sqrt{\frac{\rho}{m}} \sum_{i=1}^{m} V_{I_i, L_i} - \sqrt{m\rho} E^*[V_{I_1, L_1}] \right) + \sqrt{m\rho} E^*[V_{I_1, L_1}]. \qquad (3.9)$$

The first term of (3.9) is less than or equal to

$$\left| \sum_{i=1}^{m} V_{I_i,L_i} \right| \left| \frac{1}{\sqrt{n}} - \sqrt{\frac{\rho}{m}} \right| = \left| \frac{1}{m} \sum_{i=1}^{m} V_{I_i,L_i} \right| \left| \sqrt{m} - \sqrt{n\rho} \right| \sqrt{\frac{m}{n}} = o_p\left(\frac{1}{\sqrt{m}} \right) \xrightarrow{p^*} 0,$$

where the last equality holds by (3.8); $E^* V_{I_1,L_1} = O_p(1/\sqrt{n}\rho)$. The second term of (3.9),

$$\sqrt{m\rho}\left(\frac{1}{m} \sum_{i=1}^{m} V_{I_i,L_i} - E^*[V_{I_1,L_1}] \right) = O_p\left(\sqrt{\rho} \right) \xrightarrow{p^*} 0,$$

since $\sqrt{m}((1/m) \sum_{i=1}^{m} V_{I_i,L_i} - E^*[V_{I_1,L_1}]) = O_p(1)$ for *i.i.d.* sequence $\{V_{I_i,L_i} : i = 1, 2, \ldots\}$.

For the third term of (3.9), we have

$$E^*[V_{I_1,L_1}] = E^*[E^*(V_{I_1,L_1}|L_1)] = E^*\left[L_1\left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i \right) \right] = \frac{1}{\rho}\left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i \right).$$

Thus the third term of (3.9) is equal to

$$\frac{\sqrt{m\rho}}{n\rho} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i = \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_i'\hat{\xi}_i \right) + o_p(1) = \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{X}_i'\xi_i \right) + o_p(1) \xrightarrow{d} N(0, \Gamma).$$

Slutsky's theorem allows the desired asymptotic normality of $(1/\sqrt{n}) \sum_{i=1}^{n} \mathbf{X}_i^{*'}\hat{\xi}_i^*$ in (conditional) probability to hold and thus the proof of (a) is completed.

(b)(c) The proof of Theorem 1(b) is given by induction and by (3.1). The same arguments as in the proof of Theorem 3.1 of Thombs and Schucany (1990) are applied along with the convergence of $\hat{\phi}^*$ in conditional probability to ϕ in (3.1), and thus the detailed proof is omitted here. The proof of Theorem 1(c) is also given by the same argument as in Thombs and Schucany (1990). □

4. Concluding Remarks

This paper provides a mathematical justification for a new bootstrapping method, stationary bootstrapping, for constructing prediction intervals for GARCH models. All existing bootstrapping methods in the literature have no mathematical justifications; therefore, our mathematical results provide a theoretical contribution. It would be a salient topic for future research to make a finite sample comparison of the proposed method with other existing methods. Since the proposed method should be sensitive to the block length parameter, a detailed large sample and finite sample investigation of this issue represents an important future topic.

References

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

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001). The distribution of realized exchange rate volatility, *Journal of the American Statistical Association*, **96**, 42–55.

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

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

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

Chen, B., Gel, Y. R., Balakrishna, N. and Abraham, B. (2011). Computationally efficient bootstrap prediction intervals for returns and volatilities in ARCH and GARCH processes, *Journal of Forecasting*, **30**, 51–71.

Engle, R. F. and Patton, A. J. (2001). What good is a volatility model?, *Quantitative Finance*, **1**, 237–245.

Hwang, E. and Shin, D. W. (2011). Stationary bootstrapping for non-parametric estimator of nonlinear autoregressive model, *Journal of Time Series Analysis*, **32**, 292–303.

Hwang, E. and Shin, D. W. (2012a). Stationary bootstrap for kernel density estimators under ψ-weak dependence, *Computational Statistics and Data Analysis*, **56**, 1581–1593.

Hwang, E. and Shin, D. W. (2012b). Strong consistency of the stationary bootstrap under ψ-weak dependence, *Statistics and Probability Letters*, **82**, 488–495.

Kavalieris, L., Hannan, E. J. and Salau, M. (2003). Generalized least squares estimation of ARMA models, *Journal of Time Series Analysis*, **24**, 165–172.

Koreisha, S. and Pukkila, T. (1990). A generalized least-squares approach for estimation of autoregressive moving-average models, *Journal of Time Series Analysis*, **11**, 139–151.

Lahiri, S. N. (1999). On second-order properties of the stationary bootstrap method for studentized statistics, In *Asymptotic, Nonparametrics, and Time Series.* (Eds. Ghosh, S.), Marcel Dekker, New York, 683–711.

Miguel, J. A. and Olave, P. (1999). Bootstrapping forecast intervals in ARCH models, *Test*, **8**, 345–364.

Nordman, D. J. (2009). A note on the stationary bootstrap's variance, *Annals of Statistics*, **37**, 359–370.

Paparoditis, E. and Politis, D. N. (2005). Bootstrapping unit root tests for autoregressive time series, *Journal of the American Statistical Association*, **100**, 545–553.

Parker, C., Paparoditis, E. and Politis, D. N. (2006). Unit root testing via the stationary bootstrap, *Journal of Econometrics*, **133**, 601–638.

Pascual, L., Romo, J. and Ruiz, E. (2006). Bootstrap prediction for returns and volatilities in GARCH models, *Computational Statistics and Data Analysis*, **50**, 2293–2312.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap, *Journal of the American Statistical Association*, **89**, 1303–1313.

Reeves, J. J. (2005). Bootstrap prediction intervals for ARCH models, *Internal Journal of Forecasting*, **21**, 237–248.

Swensen, A. R. (2003). Bootstrapping unit root tests for integrated processes, *Journal of Time Series Analysis*, **24**, 99–126.

Thombs, L. A. and Schucany, W. R. (1990). Bootstrap prediction intervals for autoregression, *Journal of the American Statistical Association*, **95**, 486–492.