RESIDUAL-BASED GARCH BOOTSTRAP AND SECOND ORDER ASYMPTOTIC REFINEMENT

MINSOO JEONG
Yonsei University Wonju Campus

The residual-based bootstrap is considered one of the most reliable methods for bootstrapping generalized autoregressive conditional heteroscedasticity (GARCH) models. However, in terms of theoretical aspects, only the consistency of the bootstrap has been established, while the higher order asymptotic refinement remains unproven. For example, Corradi and Iglesias (2008) demonstrate the asymptotic refinement of the block bootstrap for GARCH models but leave the results of the residual-based bootstrap as a conjecture. To derive the second order asymptotic refinement of the residual-based GARCH bootstrap, we utilize the analysis in Andrews (2001, 2002) and establish the Edgeworth expansions of the t-statistics, as well as the convergence of their moments. As expected, we show that the bootstrap error in the coverage probabilities of the equal-tailed t-statistic and the corresponding test-inversion confidence intervals are at most of the order of $O(n^{-1})$, where the exact order depends on the moment condition of the process. This convergence rate is faster than that of the block bootstrap, as well as that of the first order asymptotic test.

1. INTRODUCTION

The residual-based bootstrap is a bootstrap method for time series models, in which the bootstrap samples are reconstructed from the estimated i.i.d. residuals. The main advantage of the residual-based bootstrap is that its convergence rate is mostly comparable to that of the i.i.d. bootstrap, whereas other generic time series bootstraps have slower convergence rates. Among generic time series bootstraps, the sieve bootstrap has a fast convergence rate equivalent to that of the i.i.d. bootstrap,[1] but it is only applicable to linear processes. While the residual-based bootstrap is applicable to a broader class of time series models, its bootstrap procedure is highly model dependent, as it actively employs the specific dependence structure of the model. In this regard, the limit theorems of the residual-based bootstraps are not readily obtainable from general bootstrap theories.

779

The generalized autoregressive conditional heteroscedasticity (GARCH) process is a good example of a time series model for which the sieve bootstrap is not applicable. First introduced by Bollerslev (1986), the GARCH model is still widely used in practice, reflecting its popularity. Hansen and Lunde (2005) investigate various volatility models in comparison with a simple GARCH(1,1) model and conclude that none of them outperforms the GARCH model when applied to exchange rate analysis. The residual-based bootstrap has also been widely used to reduce the estimation error in the GARCH model; for example, Christoffersen and Gonçalves (2005) and Mancini and Trojani (2005) use the residual-based GARCH bootstrap to obtain a more precise evaluation of the value at risk (VaR).

This paper considers the residual-based bootstrap for the GARCH(1,1) model and shows that it achieves the second order asymptotic refinement. Although there have been many papers on the residual-based GARCH bootstrap, its theoretical aspect has not been thoroughly investigated due to the complexity of the nonlinear dependence structure. Hidalgo and Zaffaroni (2007) show the consistency of the residual-based bootstrap for ARCH(∞), which encompasses GARCH(1,1). However, they do not include a result for the higher order asymptotic refinement. On the other hand, Corradi and Iglesias (2008) derive the higher order asymptotic refinement of the block bootstrap for the GARCH process, but they only conjecture that the residual-based bootstrap will provide improvement over the block bootstrap.

A bootstrap is regarded as consistent if the bootstrap test statistic successfully mimics the first order asymptotic test in the large sample. On the other hand, if the size distortion of the bootstrap test vanishes faster than that of the first order counterpart, then we say that it achieves a higher order asymptotic refinement. This higher order refinement plays an important role especially when explaining superior finite sample performances of the bootstrap. In the case of the i.i.d. bootstrap, the convergence rate of the equal-tailed t-statistic is $O\left(n^{-1}\right)$, which is faster than the first order counterpart $O\left(n^{-1/2}\right)$. However, unlike the i.i.d. bootstrap, bootstraps for time series models have different convergence rates depending on the model and applied method. The bootstrap methods applicable to the generic time series model usually present slower convergence rates than the i.i.d. bootstrap. For example, the block bootstrap, which can be applied to virtually any stationary time series, has a slower convergence rate than other bootstrap methods. The nonparametric Markov bootstrap has a faster convergence rate than the block bootstrap, as shown in Horowitz (2003). Nonetheless, its convergence rate is slower than that of the i.i.d. bootstrap, and its performance depends on the choice of bandwidths.

However, when the exact form of the dependence structure is known, a greater improvement can be achieved by using a model specific bootstrap method. For example, Bose (1988) and Bose (1990) analyze bootstrap methods for the autoregressive (AR) and moving average (MA) models, respectively. Andrews (2005) suggests a parametric Markov bootstrap, which can be applied to Markov processes with known dependence structures. Although this parametric bootstrap is applicable to a wide range of Markov processes, the GARCH process is not

Markovian.[2] Therefore, the theoretical analysis of the parametric Markov bootstrap is not applicable, and that is why a new approach is needed to analyze the residual-based GARCH bootstrap.

This paper provides higher order asymptotic analyses of the residual-based GARCH bootstrap. To this end, we first exploit the lemmas in Andrews (2001, 2002) to show that the t-statistic of the GARCH maximum likelihood (ML) estimator admits an Edgeworth expansion, whose coefficients are determined by the moments of the statistic. By establishing the convergence of the bootstrap moments to the population moments, we then show that the bootstrap error in the coverage probability (ECP) of the equal-tailed t-statistic is of the order of $o(n^{-1+\delta})$ for some $\delta > 0$ given by the moment condition. This convergence rate is faster than that of the block bootstrap, as well as that of the first order asymptotics. We also construct bootstrap confidence intervals using the test-inversion technique and show that our bootstrap confidence intervals share the faster convergence rate of the bootstrap t-test. The finite sample properties of the residual-based GARCH bootstrap have been studied in various papers, including Pascual et al. (2006), Chen et al. (2011), and Varga and Zempléni (2012). They assure that the residual-based GARCH bootstrap provides nontrivial finite sample refinements over the first order asymptotics. Therefore, we omit the finite sample analysis of the GARCH bootstrap in this paper.

The remainder of the paper is organized as follows. Section 2 defines the model and provides a detailed description of the residual-based bootstrap for the GARCH model. Section 3 states the assumptions and main asymptotic result of the paper. Section 4 provides the research conclusion.

2. GARCH BOOTSTRAP

2.1. The Model

We consider the GARCH(1,1) process $\{u_t\}$ defined by

$$
\begin{aligned}
u_t &= \sigma_t \varepsilon_t, \\
\sigma_t^2 &= \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2,
\end{aligned}
\tag{1}
$$

where $\omega, \alpha, \beta > 0$, and $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with zero mean and unit variance. The GARCH process $\{u_t\}$ defined in (1) becomes a zero mean white noise process, whose conditional variance is given by $\mathbb{E}(u_t^2|\mathcal{F}_{t-1}) = \sigma_t^2$. The squared return process $\{u_t^2\}$ follows the ARMA(1,1) model with heteroscedastic errors such that

$$
u_t^2 = \omega + (\alpha + \beta) u_{t-1}^2 + v_t - \beta v_{t-1},
$$

where $v_t = u_t^2 - \sigma_t^2$. From this representation, we often interpret $\alpha + \beta$ as the persistency parameter of the volatility, and the model (1) is called the integrated GARCH (IGARCH) when $\alpha + \beta = 1$.

The stationarity of $\{u_t\}$ is determined by the distribution of $\{\varepsilon_t\}$ and the parameter values of α and β. More precisely, the strictly stationary condition of the GARCH process is obtained in Nelson (1990) and Klüppelberg et al. (2004) as

$$-\infty < \mathbb{E}\log\left(\alpha\varepsilon_t^2 + \beta\right) < 0. \tag{2}$$

Note that (2) implies $\beta < 1$. In this paper, we only consider strictly stationary GARCH processes and let σ_0^2 start from the stationary distribution of $\{\sigma_t^2\}$, which is given by

$$\omega\left[1 + \sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\alpha\varepsilon_{-i}^2 + \beta\right)\right]. \tag{3}$$

For the IGARCH model, there exists a strictly stationary solution for the IGARCH(1,1) model satisfying (2).[3] However, our result does not include the IGARCH model since our conditions in Section 3.1 require the existence of the second moment, while the IGARCH model has an infinite second moment.

We consider the quasi-maximum likelihood (QML) estimator of the GARCH model, which is given by the maximizer of $\sum_{t=1}^{n}\ell_t(\theta)$ over θ in the interior of Θ, where $\theta = (\omega, \alpha, \beta)'$, Θ is the parameter space, and

$$\ell_t = -\frac{1}{2}\log\sigma_t^2 - \frac{1}{2}\frac{u_t^2}{\sigma_t^2},$$

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta\sigma_{t-1}^2 = \omega\sum_{i=0}^{t-1}\beta^i + \alpha\sum_{i=0}^{t-1}\beta^i u_{t-1-i}^2 + \beta^t\sigma_0^2. \tag{4}$$

As noted earlier, the GARCH process is not Markovian, and the likelihood ℓ_t of the GARCH QML estimator depends on the whole history of $\{u_t\}$, distinguishing it from the usual Markov ML estimators. The consistency and asymptotic normality of this QML estimator were first shown in Lee and Hansen (1994) and Lumsdaine (1996) and later refined with much weaker conditions in Berkes et al. (2003) and Francq and Zakoïan (2004). The former establishes the asymptotic normality of the estimator under minimal conditions on the innovations $\{\varepsilon_t\}$, and the latter further reduces technical assumptions on the distribution of the innovations at the origin. In the following sections, we heavily rely on their results of the QML estimator properties to derive our asymptotic results. On the other hand, the case in which the distribution of the innovation $\{\varepsilon_t\}$ is heavy-tailed is also an important and interesting subject. Berkes and Horváth (2003) show asymptotic properties of QML estimators in the presence of heavy-tails in $\{\varepsilon_t\}$, and Hall and Yao (2003) consider the case of infinite variance such that $\mathbb{E}\varepsilon_t^2 = \infty$. The rate of convergence and sometimes even the limit distributions for the case of heavy-tailed innovations differ from those of the usual stationary asymptotics, and we do not cover this case in the present paper.

For the test statistic in this paper, we define the t-statistic of the QML estimator testing the null hypothesis $H_0 : \theta_r = \theta_{0,r}$ as

$$T_n = n^{1/2} \frac{\hat{\theta}_r - \theta_{0,r}}{\hat{\varsigma}_{rr}^{1/2}},$$

where $\hat{\theta}_r$ is the rth element of $\hat{\theta}$, and $\hat{\varsigma}_{rr}$ is the (r, r) component of the variance matrix estimator $\hat{\varsigma}$. We use the variance matrix estimator $\hat{\varsigma}$ given by

$$\hat{\varsigma} = n \left(\sum_{i=1}^{n} \frac{\partial^2 \ell_i(\hat{\theta})}{\partial \theta \partial \theta'} \right)^{-1} \sum_{i=1}^{n} \frac{\partial \ell_i(\hat{\theta})}{\partial \theta} \frac{\partial \ell_i(\hat{\theta})}{\partial \theta'} \left(\sum_{i=1}^{n} \frac{\partial^2 \ell_i(\hat{\theta})}{\partial \theta \partial \theta'} \right)^{-1}.$$

2.2. Residual-Based Bootstrap

This section describes the residual-based bootstrap procedure, which is specifically designed for the GARCH process in (1). This is a natural extension of the residual-based bootstrap for AR models and is well illustrated in Pascual et al. (2006) and Shimizu (2009), as well as in other literature.

The detailed procedure reads as follows. (i) Estimate $\hat{\omega}$, $\hat{\alpha}$, and $\hat{\beta}$ using the QML estimation with $\{u_t\}$. (ii) Construct $\{\hat{\varepsilon}_t\}$ using the following formulae recursively:

$$\begin{aligned}
\hat{\sigma}_t^2 &= \hat{\omega} + \hat{\alpha} u_{t-1}^2 + \hat{\beta} \hat{\sigma}_{t-1}^2, \\
\hat{\varepsilon}_t &= u_t / \hat{\sigma}_t,
\end{aligned} \tag{5}$$

where $u_0 = \hat{\sigma}_0 \hat{\varepsilon}_0$, and $\hat{\sigma}_0^2$ and $\hat{\varepsilon}_0$ are obtained from the stationary distribution of $\{\hat{\sigma}_t^2\}$ and the distribution of $\{\hat{\varepsilon}_t\}$, respectively, for the given parameter values of $\hat{\omega}$, $\hat{\alpha}$, and $\hat{\beta}$.[4] (iii) Bootstrap $\{\varepsilon_t^*\}$ from the empirical distribution of demeaned $\{\hat{\varepsilon}_t\}$. (iv) Construct $\{u_t^*\}$ using the bootstrapped innovations $\{\varepsilon_t^*\}$ in (iii). (v) Estimate $\hat{\omega}^*$, $\hat{\alpha}^*$, and $\hat{\beta}^*$ using the bootstrapped sample $\{u_t^*\}$ and repeat (iii)–(v) to obtain more bootstrap estimates. Note here that, in (iii), we acquire the estimated innovations $\{\varepsilon_t^*\}$, which are independent and identically distributed conditional on the sample $\{u_t\}$.

3. ASYMPTOTIC REFINEMENT

3.1. Regularity Conditions

We define the following notations to be used throughout the paper. Let

$$s_t = \frac{\partial \ell_t}{\partial \theta} = \frac{u_t^2}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} - \frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} \tag{6}$$

and φ_t be vectors containing the unique components of s_t and its derivatives through order 6 with respect to θ. We also denote $(\partial^j / \partial \theta^j) s_t(\theta)$ and

$(\partial^j / \partial \theta^j) \varphi_t(\theta)$ as the vectorized partial derivatives with respect to θ of order j. For each element of φ_t, we let $\varphi_{t,k}$ be the kth element of φ_t. We let F be the distribution function of $\{\varepsilon_t\}$. We similarly define the bootstrap version of the notations using a superscript '$*$' such that F^* denotes the distribution function of $\{\hat{\varepsilon}_t\}$, for example.

We assume the following assumptions to derive the asymptotic properties of the bootstrap.

Assumption 1.

  (a) Θ is compact, and all $\theta \in \Theta$ satisfy $\omega, \alpha, \beta > 0$ and $\mathbb{E}(\alpha \varepsilon_t^2 + \beta)^{p/2} < 1$ for some $p > 6$.
  (b) θ_0 is an interior point of Θ.
  (c) $F(-x)$ and $1 - F(x)$ exponentially decay as $x \to \infty$.
  (d) F is continuously differentiable and globally Lipschitz continuous.

We require Assumption 1(a) mainly to guarantee the moment conditions set out by Andrews (2001). Assumption 1(b) is to avoid the boundary value problem of the bootstrap, as illustrated in, e.g., Andrews (2000). Assumptions 1(c) and 1(d) are to ensure the β-mixing property of the process and obtain a proper convergence rate of the bootstrap moments of the test statistics. It is also notable that Assumptions 1(a) and 1(d) lead to the stationarity condition (2).

We follow steps similar to those in Andrews (2001) to derive our asymptotic results in this paper. Therefore, we need to ensure that the conditions in Andrews (2001) are satisfied. For example, Assumptions 1(b) and 1(c) guarantee the negative definiteness of the Hessian,[5] and Assumption 1(a) is a sufficient condition for the existence of the 6th moment,[6] which ensures the moment conditions on the derivatives of the log-likelihood in Andrews (2001).[7] We only need $p > 0$ to satisfy the conditions in Andrews (2001), but we further require $p > 6$ in this paper to address initial value effects of the residual-based bootstrap.

Assumption 1(c) is also necessary to ensure a dependent version of the Cramér condition: there exist $\Lambda < \infty$ and $\upsilon > 0$ such that, for arbitrary large $\zeta > 1$ and all integers $m \in (\upsilon^{-1}, N)$ and $\tau \in \mathbb{R}^{\dim(\varphi_s)}$ with $\upsilon < \|\tau\| < N^{\zeta}$,

$$\mathbb{E} \left| \mathbb{E} \left(\exp \left[\sqrt{-1}\, \tau' \sum_{s=1}^{2m+1} \varphi_s(\theta_0) \right] \middle| \{\varepsilon_j : |j - m| > \Lambda\} \right) \right| \leq \exp(-\upsilon), \qquad (7)$$

where θ_0 is the true parameter of the model. This Cramér condition is necessary for the existence of the Edgeworth expansion of the normalized sample mean. It is shown in Lemma A.1 of Corradi and Iglesias (2008) that (7) holds under our set of assumptions.

Furthermore, Assumption 1(c) ensures the β-mixing properties of $\{\varphi_t\}$, which is essential to derive the convergence of the bootstrap moments. It is shown in Francq and Zakoïan (2006) that a general class of GARCH processes is β-mixing with an exponentially decaying coefficient under mild assumptions. It is also

shown in Corradi and Iglesias (2008) that $\{\varphi_t\}$, which is essentially a vector of derivatives of the GARCH process with respect to the parameters, is β-mixing with an exponentially decaying rate.

In addition to these conditions, Andrews (2001) also requires that there exist $K < \infty$ and $c > 0$ such that, for $m \geq 1$,

$$\mathbb{E}\big|\eta_k(\varepsilon_t, \ldots, \varepsilon_0) - \eta_k(\varepsilon_t, \ldots, \varepsilon_{t-m}, 0, \ldots, 0)\big| \leq K \exp(-cm), \tag{8}$$

where η_k is defined such that $\varphi_{t,k} = \eta_k(\varepsilon_t, \ldots, \varepsilon_0)$. The β-mixing properties of $\{\varphi_t\}$ can be used here, and the proof of Theorem 1 of Corradi and Iglesias (2008) shows that (8) holds under Assumptions 1(b) and 1(c).

3.2. Second Order Refinement

In this section, we first show in Theorem 1 that the error in the coverage probabilities of the bootstrap t-tests has a fast convergence rate. Using this limit theorem, we then suggest a bootstrap test-inversion method to obtain refined confidence intervals in Corollary 2.

THEOREM 1. *Let Assumption 1 hold. Then, under the null hypothesis, we have*

$$\mathbf{P}\left(T_n \geq z^*_{a/2} \ \text{or} \ T_n \leq z^*_{1-a/2}\right) = a + o\left(n^{-1+\delta}\right)$$

as $n \to \infty$ for any $\delta > 2/p$. The bootstrap critical value z^*_a is defined as the $1-a$ quantile of T^*_n, where $T^*_n = n^{1/2}(\hat{\theta}^*_r - \hat{\theta}_r)/\hat{\varsigma}^{*1/2}_{rr}$.

The convergence rate of the first order asymptotic approximation under the null hypothesis is given by

$$\mathbf{P}\left(T_n \geq z_{a/2} \ \text{or} \ T_n \leq -z_{a/2}\right) = a + O\left(n^{-1/2}\right)$$

as $n \to \infty$, where z_a denotes the $1-a$ quantile of the standard normal distribution $\mathbb{N}(0, 1)$. Therefore, the theorem shows that the residual-based bootstrap reduces the error in the coverage probability by up to $O\left(n^{-1/2}\right)$.

Compared with the convergence rate of the block bootstrap, which is derived in Corradi and Iglesias (2008) as $o\left(n^{-1/2-\xi}\right)$ for some $0 \leq \xi < 1/4$, our convergence rate $o\left(n^{-1+\delta}\right)$ is always faster as long as $p \geq 8$ since $\delta > 2/p$. If we consider the ideal situations both for the residual-based and block bootstraps, then the difference in their convergence rates is of the order of $n^{1/4}$. Moreover, while ξ for the block bootstrap is determined by the choice of the block length parameter, δ is only given by the moment condition; thus, we can make δ arbitrarily small as long as the tail of $\{u_t\}$ is decreasing fast enough. Therefore, we may say that the performance of the residual-based bootstrap is essentially of the order of $O\left(n^{-1}\right)$ given a sufficient moment condition, while the performance of the block bootstrap varies across the choices of block length and other sample properties of the data.

A good example as an application of Theorem 1 is to obtain confidence intervals of a high precision. For hypothesis testing of the GARCH model, it is often of interest whether or not the data has conditional heteroscedasticity, which naturally involves a null hypothesis of zero parameter values. However, our GARCH bootstrap cannot deal with the boundary parameter value cases, as noted in Andrews (2000). Therefore, the asymptotic result in Theorem 1 may not be of great interest to practitioners. However, utilizing the duality between the hypothesis testing and confidence interval estimation, we can obtain a refined confidence interval of the estimator directly from the bootstrap t-test critical values. The $100(1-a)\%$ confidence interval obtained from the test-inversion technique is given by $[\kappa_L, \kappa_U]$, where κ_L and κ_U are solutions to

$$
\begin{aligned}
\mathbf{P}^* \left(\frac{\hat{\theta}_r^* (\hat{\theta}^{\kappa_U}) - \kappa_U}{\hat{\varsigma}_{rr}^{*1/2} (\hat{\theta}^{\kappa_U})} \leq \frac{\hat{\theta}_r - \kappa_U}{\hat{\varsigma}_{rr}^{1/2}} \right) &= \frac{a}{2}, \\
\mathbf{P}^* \left(\frac{\hat{\theta}_r^* (\hat{\theta}^{\kappa_L}) - \kappa_L}{\hat{\varsigma}_{rr}^{*1/2} (\hat{\theta}^{\kappa_L})} \geq \frac{\hat{\theta}_r - \kappa_L}{\hat{\varsigma}_{rr}^{1/2}} \right) &= \frac{a}{2},
\end{aligned}
\tag{9}
$$

where $\hat{\theta}_r^* (\hat{\theta}^{\kappa_U})$ is the bootstrap version of the estimator of θ_r obtained from a sample generated at parameter values $\hat{\theta}^{\kappa_U}$, and $\hat{\varsigma}_{rr}^* (\hat{\theta}^{\kappa_U})$ is the corresponding variance estimator. Here we denote $\hat{\theta}^{\kappa_U}$ as the parameter estimator $\hat{\theta}$, whose rth element is replaced by κ_U. We define $\hat{\theta}_r^* (\hat{\theta}^{\kappa_L})$ and $\hat{\varsigma}_{rr}^* (\hat{\theta}^{\kappa_L})$ in the same way as $\hat{\theta}_r^* (\hat{\theta}^{\kappa_U})$ and $\hat{\varsigma}_{rr}^* (\hat{\theta}^{\kappa_U})$, respectively. This is called the standardized test-inversion bootstrap (STIB) confidence interval. This test-inversion technique is introduced in Carpenter (1999) and many others, with slightly different definitions.

For this STIB confidence interval, we obtain the following corollary from Theorem 1.

COROLLARY 2. *Let κ_U and κ_L satisfy (9) for some $a \in (0,1)$. Then, under Assumption 1, we have*

$$
\mathbf{P} \left(\theta_{0,r} \geq \kappa_U \text{ or } \theta_{0,r} \leq \kappa_L \right) = a + o \left(n^{-1+\delta} \right)
$$

as $n \to \infty$ for any $\delta > 2/p$, where θ_0 is the true parameter value of the model.

The convergence rate of the first order asymptotic approximation is given by

$$
\mathbf{P} \left(\theta_{0,r} \geq \hat{\theta}_r + z_{a/2} \frac{\hat{\varsigma}_{rr}^{1/2}}{n^{1/2}} \text{ or } \theta_{0,r} \leq \hat{\theta}_r - z_{a/2} \frac{\hat{\varsigma}_{rr}^{1/2}}{n^{1/2}} \right) = a + O \left(n^{-1/2} \right)
$$

as $n \to \infty$. Therefore, the confidence interval obtained from the residual-based bootstrap also reduces the error in the coverage probability by up to $O \left(n^{-1/2} \right)$.

4. CONCLUSION

In this paper, we revisit the residual-based bootstrap for the GARCH(1,1) model and show that it achieves asymptotic refinement. Utilizing the analysis of Andrews (2001, 2002), we show that it has a faster convergence rate than the first order limit distributions, under mild assumptions on the tail behaviors and smoothness of the distribution function of the innovations. The convergence order of the residual-based bootstrap is obtained as $o\left(n^{-1+\delta}\right)$ for some $\delta > 0$ given by the moment condition, which is faster than $O\left(n^{-1/2}\right)$, the convergence rate of the first order limit distribution.

Though the block bootstrap can ideally achieve the convergence rate $O\left(n^{-3/4}\right)$ by appropriately choosing the block length parameter, the performance of the block bootstrap heavily depends on the block size, which is often not easy to determine optimally from a given sample. Unlike the block bootstrap, the residual-based bootstrap is robust to user dependent choices and achieves a faster second order asymptotic refinement comparable to that of the i.i.d. bootstrap, as long as sufficient moment conditions are satisfied.

NOTES

1. See Choi and Hall (2000).
2. The GARCH process is Markovian if we consider the bivariate process $\{u_t, \sigma_t^2\}$ defined in (1).
3. The IGARCH(1,1) model has an infinite second moment and, therefore, is not covariance stationary.
4. The stationary distribution of $\{\hat{\sigma}_t^2\}$ in (3) is generally not given in a closed-form. Therefore, we rely on numerical methods to obtain the initial values of the bootstrap. For example, first set arbitrary values for $\hat{\sigma}_0^2$ and $\hat{\varepsilon}_0$. Then we obtain the distribution of $\{\hat{\varepsilon}_t\}$ by generating $\hat{\varepsilon}_t$ for $t = 1, \ldots, n$ using (5). Next, we randomly draw $\hat{\varepsilon}_0$ and $\hat{\sigma}_0^2$ using the distribution of $\{\hat{\varepsilon}_t\}$ and the formula

$$\hat{\sigma}_0^2 = \hat{\omega}\left[1 + \sum_{k=1}^{m}\prod_{i=1}^{k}\left(\hat{\alpha}\hat{\varepsilon}_{-i}^2 + \hat{\beta}\right)\right] \tag{10}$$

for some large $m > 0$, where m is chosen for the desired precision. We iterate (5) and (10), then the initial values for (ii) are obtained up to an arbitrary precision.
5. See Theorem 4.2 and Remark 4.5 of Berkes et al. (2003).
6. See, e.g., Davis and Mikosch (2009).
7. See the proof of Lemma 1 in the online supplement for details.

REFERENCES

Andrews, D.W.K. (2000) Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica* 68, 399–405.

Andrews, D.W.K. (2001) Higher-order improvements of a computationally attractive k-step bootstrap for extremum estimators. Cowles Foundation Discussion Paper No. 1230R, Cowles Foundation for Research in Economics, Yale University.

Andrews, D.W.K. (2002) Higher-order improvements of a computationally attractive k-step bootstrap for extremum estimators. *Econometrica* 70, 119–162.

Andrews, D.W.K. (2005) Higher-order improvements of the parametric bootstrap for Markov processes, in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge, pp. 171–215.

Berkes, I., & L. Horváth (2003) The rate of consistency of the quasi-maximum likelihood estimator. *Statistics & Probability Letters* 61, 133–143.

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

Bhattacharya, R.N. (1987) Some aspects of Edgeworth expansions in statistics and probability, in M. L. Puri, J. P. Vilaplana, and Wolfgang Wertz (eds.), *New Perspectives in Theoretical and Applied Statistics*, John Wiley & Sons, pp. 157–170.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics* 31(3), 307–327.

Bose, A. (1988) Egdeworth correction by bootstrap in autoregressions. *Annals of Statistics* 16, 1709–1722.

Bose, A. (1990) Bootstrap in moving average models. *Annals of the Institute of Statistical Mathematics* 42, 753–768.

Carpenter, J. (1999) Test inversion bootstrap confidence intervals. *Journal of the Royal Statistical Society: Series B* 61(1), 159–172.

Chen, B., Y.R. Gel, N. Balakrishna, & B. Abraham (2011) Computationally efficient bootstrap prediction intervals for returns and volatilities in ARCH and GARCH processes. *Journal of Forecasting* 30(1), 51–71.

Choi, E., & P. Hall (2000) Bootstrap confidence regions computed from autoregressions of arbitrary order. *Journal of the Royal Statistical Society, Series B* 62, 461–477.

Christoffersen, P., & S. Gonçalves (2005) Estimation risk in financial risk management. *Journal of Risk* 7, 1–29.

Corradi, V., & E. Iglesias (2008) Bootstrap refinements for QML estimators of the GARCH(1,1) parameters. *Journal of Econometrics* 144(2), 500–510.

Davis, R.A., & T. Mikosch (2009) Extreme value theory for GARCH processes. in *Handbook of Financial Time Series*, Springer, pp. 187–200.

Doukhan, P. (1994) *Mixing: Properties and Examples*, New York: Springer-Verlag.

Francq, C., & J. Zakoïan (2004) Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* 10(4), 605–637.

Francq, C., & J. Zakoïan (2006) Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory* 22(5), 815–834.

Götze, F., & C. Hipp (1994) Asymptotic distribution of statistics in time series. *Annals of Statistics* 22, 2062–2088.

Hall, P., & Q. Yao (2003) Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* 71(1), 285–317.

Hansen, P.R., & A. Lunde (2005) A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics* 20(7), 873–889.

Hidalgo, J., & P. Zaffaroni (2007) A goodness-of-fit test for ARCH(∞) models. *Journal of econometrics* 141(2), 835–875.

Horowitz, J.L. (2003) Bootstrap methods for Markov processes. *Econometrica* 71, 1049–1082.

Klüppelberg, C., A. Lindner, & R. Maller (2004) A continuous-time GARCH process driven by a Levy process: stationary and second order behaviour. *Journal of Applied Probability* 41, 601–622.

Lee, S.W., & B. E. Hansen (1994) Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory* 10, 29–29.

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(3), 575–596.

Mancini, L., & F. Trojani (2005) Robust semiparametric bootstrap methods for value at risk prediction under GARCH-type volatility processes. Working Paper, University of St. Gallen.

Mikosch, T., & C. Stărică (2000) Limit theory for the sample autocorrelations and extremes of a GARCH (1,1) process. *The Annals of Statistics* 28(5), 1427–1451.

Nelsen, D.B. (1990) Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* 6, 318–334.

Pascual, L., J. Romo, & E. Ruiz (2006) Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics and Data Analysis* 50(9), 2293–2312.

Rio, E. (1993) Covariance inequalities for strongly mixing processes. *Annales de l'Institut Henri Poincaré (B)* 29(4), 587–597.

Shimizu, K. (2009) *Bootstrapping stationary ARMA-GARCH models*. Vieweg+Teubner.

Varga, L., & A. Zempléni (2012) Weighted bootstrap in GARCH models. Working Paper, Eötvös Loránd University.

APPENDIX A

In this section, we provide useful lemmas and the proofs of theorems. Here we only present the outline of the proofs and leave the rest of them to the online supplement.

A.1. Useful Lemmas

LEMMA A.1. *We have*

$$\lim_{n\to\infty} n\mathbf{P}\left(n^{1/2-\delta}\sup_{x\in\mathbb{R}}\left|F^*(x)-F(x)\right|>\epsilon\right)=0$$

for any $\epsilon>0$ and $\delta>2/p$.

Proof. To show the statement of the lemma, we first show the uniform convergence between $\hat{\varepsilon}_t$ and ε_t in the sense that

$$\lim_{n\to\infty} n\mathbf{P}\left(n^{1/2-\delta}|\hat{\varepsilon}_t-\varepsilon_t|>\epsilon\right)=0 \tag{A.1}$$

uniformly in $n^a\leq t\leq n$ for arbitrarily small $a>0$, and then establish the convergence between their distribution functions F^* and F using (A.1) and the convergence theorem for extremal processes in Mikosch and Stărică (2000). We omit the proof here and provide detailed derivations in the online supplement. ∎

LEMMA A.2. *Let $S_n=n^{-1}\sum_{t=1}^n\varphi_t(\theta_0)$, $S=\mathbb{E}S_n$, $S_n^*=n^{-1}\sum_{t=1}^n\varphi_t^*(\hat{\theta})$ and $S^*=\mathbb{E}^*S_n^*$. Further define $\Psi_n=n^{1/2}(S_n-S)$ and $\Psi_n^*=n^{1/2}(S_n^*-S^*)$, and denote $\Psi_{n,k}$ and $\Psi_{n,k}^*$ as the kth elements of Ψ_n and Ψ_n^*, respectively. Also, let v_n and v_n^* be vectors of moments of the form $n^{\alpha(m)}\mathbb{E}\prod_{\mu=1}^m\Psi_{n,k_\mu}$ and $n^{\alpha(m)}\mathbb{E}^*\prod_{\mu=1}^m\Psi_{n,k_\mu}^*$ for $k_\mu\in\{1,\ldots,\dim(\Psi_n)\}$, respectively, where $\alpha(m)=0$ when m is even and $\alpha(m)=1/2$ when m is odd, for $2\leq m\leq 4$, and let $v=\lim_{n\to\infty}v_n$. Then we have*

$$\lim_{n\to\infty} n\mathbf{P}\left(n^{1/2-\delta}\left\|v_n^*-v\right\|>\epsilon\right)=0$$

for any $\epsilon>0$ and $\delta>2/p$.

Proof. Let $\psi_t=\prod_{i=1}^k[\varphi_{t,r_i}(\theta_0)-\mathbb{E}\varphi_{t,r_i}(\theta_0)]$ and $\psi_t^*=\prod_{i=1}^k\left[\varphi_{t,r_i}^*(\hat{\theta})-\mathbb{E}^*\varphi_{t,r_i}^*(\hat{\theta})\right]$ for $k>0$ and $r_i\in\{1,\ldots,\dim(\varphi_t)\}$. To prove the statement of the lemma, we first focus on the convergence of $\mathbb{E}^*(\psi_t^*)-\mathbb{E}(\psi_t)$, and show that each additive term of $\varphi_{t,r}(\theta_0)-\mathbb{E}\varphi_{t,r}(\theta_0)$ as a function of ε_i for $i=0,\ldots,t$ has a specific functional form,

where we omit the details here. Then utilizing this specific functional form and the convergence of the distribution function in Lemma A.1, we deduce that

$$\mathbb{E}^*\left(\psi_t^*\right) - \mathbb{E}(\psi_t) = O\left(n^{-1/2+\delta}\right) + O\left(\eta^t\right) \tag{A.2}$$

for some $0 < \eta < 1$ as $n \to \infty$ and $t \to \infty$ with probability $1 - o(n^{-1})$, where the detailed derivations are omitted here. After we establish the convergence order in (A.2), we obtain the main statement of the lemma as follows:

$$n^{\alpha(m)}\mathbb{E}^* \prod_{\mu=1}^{m} \Psi_{n,k_\mu}^* - \lim_{n\to\infty} n^{\alpha(m)}\mathbb{E} \prod_{\mu=1}^{m} \Psi_{n,k_\mu} = O\left(n^{-1/2+\delta}\right)$$

as $n \to \infty$ for $m = 2, 3, 4$ with probability $1 - o(n^{-1})$, where the convergence order is obtained similar to the proofs in Andrews (2002), using the covariance inequality in Rio (1993), the strong mixing inequality in Doukhan (1994), and the β-mixing properties of $\{\varphi_t\}$. We also omit the derailed derivation here and leave it to the online supplement. ∎

A.2. Proofs of Theorem 1 and Corollary 2

The proofs for Theorem 1 and Corollary 2 are slight modifications of the proofs in Andrews (2002) and Carpenter (1999), respectively, for which we utilize the convergence of the moments in Lemma A.2, as well as the results in Bhattacharya (1987) and Götze and Hipp (1994). We omit the proof here, and provide detailed derivations in the online supplement.