

Bootstrap-Based Test for Volatility Shifts in GARCH against Long-Range Dependence

Yu Wang^a, Cheolwoo Park^a, Taewook Lee^{1,b}

^aDepartment of Statistics, University of Georgia, USA;

^bDepartment of Statistics, Hankuk University of Foreign Studies, Korea

Abstract

Volatility is a variation measure in finance for returns of a financial instrument over time. GARCH models have been a popular tool to analyze volatility of financial time series data since Bollerslev (1986) and it is said that volatility is highly persistent when the sum of the estimated coefficients of the squared lagged returns and the lagged conditional variance terms in GARCH models is close to 1. Regarding persistence, numerous methods have been proposed to test if such persistency is due to volatility shifts in the market or natural fluctuation explained by stationary long-range dependence (LRD). Recently, Lee *et al.* (2015) proposed a residual-based cumulative sum (CUSUM) test statistic to test volatility shifts in GARCH models against LRD. We propose a bootstrap-based approach for the residual-based test and compare the sizes and powers of our bootstrap-based CUSUM test with the one in Lee *et al.* (2015) through simulation studies.

Keywords: Bootstrap, CUSUM test, GARCH models, Long-range dependence, Volatility shifts

1. Introduction

In finance, volatility is a variation measure for returns of a financial instrument over time. Measuring volatility has been one of the most critical steps in financial modeling. A volatility model can be used to forecast the absolute magnitude of returns or the density of the volatility in risk management, asset pricing, and portfolio selection. Engle (1982) introduced the autoregressive conditional heteroscedastic (ARCH) process and mentioned several attractive characteristics of ARCH models. A more general process, generalized autoregressive conditional heteroscedasticity (GARCH) models, have been a popular tool to analyze volatility of financial time series data since Bollerslev (1986). GARCH (p, q) models satisfy:

$$x_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega^\circ + \sum_{i=1}^q \alpha_i^\circ x_{t-i}^2 + \sum_{j=1}^p \beta_j^\circ \sigma_{t-j}^2, \quad (1.1)$$

where the innovation $\{\epsilon_t\}_{t \in \mathbb{Z}}$ is a sequence of standard i.i.d. random variables. It is assumed that $\omega^\circ > 0$, $\alpha_i^\circ \geq 0$ for all $i = 1, \dots, q$ and $\beta_j^\circ \geq 0$ for all $j = 1, \dots, p$. It is said that the volatility is highly persistent when the sum of the estimated coefficients of the squared lagged returns and the lagged conditional variance terms in GARCH models is close to 1.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2013R1A1A2006796).

¹ Corresponding author: Department of Statistics, Hankuk University of Foreign Studies, Yongin 17035, Korea.
E-mail: twlee@hufs.ac.kr

Published 30 September 2015 / journal homepage: <http://csam.or.kr>

© 2015 The Korean Statistical Society, and Korean International Statistical Society. All rights reserved.

An alternative way to deal with the high persistence of volatility is to use long-range dependent (LRD) modeling on transformed data such as squared returns. Beran (1994) pointed out that scientists in various areas have observed correlations between observations that decay to zero at a slow rate. This long memory phenomenon has been documented long before suitable models were developed. Ding *et al.* (1993) investigated the long-memory property of the stock market return series. They found that the power transformation of the absolute returns has quite high autocorrelations for long lags. A LRD time series $\{x_t\}_{t \in \mathbb{Z}}$ is defined as a second-order stationary time series model with a slowly decaying autocovariance function,

$$\gamma(h) = \text{Cov}(x_t, x_{t+h}) \sim Ch^{2d-1} = Ch^{2H-2}, \quad \text{as } h \rightarrow \infty, \quad (1.2)$$

where $C > 0$ is a constant,

$$d \in \left(0, \frac{1}{2}\right)$$

is the long-range dependence parameter, and

$$H = d + \frac{1}{2} \in \left(\frac{1}{2}, 1\right)$$

is the self-similarity (SS) parameter. Based on (1.2) for LRD series, the autocovariances are not absolute summable. That is, $\sum_{h=-\infty}^{\infty} |\gamma(h)| = +\infty$. If, on the other hand, the autocovariances are absolute summable, $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, the series is short-range dependence (SRD).

First proposed by Efron (1979), the bootstrap method is another key tool in this paper. Bootstrap has been developed to overcome inaccuracies caused by small sample sizes in statistical inference. There are many different types of bootstrap methods, but not all of them are easy to implement or work well in certain cases. In this paper, we use the standard residual-based bootstrap method with the procedures presented in Section 3.

This paper focuses on the persistence in volatility of financial time series. Numerous methods have been proposed to test if such persistency is due to volatility shifts in the market or natural fluctuation explained by stationary long-range dependence. Lee *et al.* (2015) proposed a residual-based cumulative sum (CUSUM) test statistic to test volatility shifts in GARCH models against LRD. In this paper, we propose a bootstrap-based approach for the residual-based test and compare the sizes and powers of our bootstrap-based CUSUM test with the one in Lee *et al.* (2015) through simulation studies.

The paper is organized as follows. Section 2 introduces the tests on volatility shift versus LRD in Lee *et al.* (2015). Section 3 proposes the bootstrap-based CUSUM method for the residual-based test. Simulation results are reported in Section 4 and the conclusion in Section 5.

2. Residual-Based CUSUM Test for Volatility Shifts against LRD

Many statistical tests to distinguish LRD and volatility shift divide the procedure into two classes according to the null hypotheses: either LRD as the null hypothesis (Qu, 2011; Kuswanto, 2011), or the volatility-shift model as the null (Jach and Kokoszka, 2008). In this paper, we use the latter one as the null hypothesis. The tests consider:

$$H_0^{(R)} : \text{VS-R model} \quad \text{versus} \quad H_1 : \text{LRD model},$$

where VS-R represents that volatility shifts occurred at R unknown locations. We fit each of the $(R+1)$ regimes using GARCH models.

Researchers proposed various methods to estimate the unknown change locations for the volatility. The most popular is the CUSUM type of test statistic first considered by Kokoszka and Leipus (2000), where they use the squared asset returns for a single volatility shift. Andreou and Ghysels (2002) argue that this test suffers from size distortions under the finite sample because of the strong dependence on observations. The Bartlett long-run variance estimator can be used to correct the size distortion (Berkes *et al.*, 2006; Zhang *et al.*, 2007). However, the Bartlett long-run variance estimator is sensitive to choice of the kernel bandwidth. To overcome such defects, Lee *et al.* (2015) adopted the residual-based CUSUM test. The test uses the standardized residuals from an estimated GARCH model that imitate the innovation series in GARCH models and reduce the dependence among observations. Unlike the statistic proposed by Lee *et al.* (2004), the residual-based CUSUM test statistic does not require tuning parameters for truncation and can be used for various types of volatility models.

The power transformation of returns, such $\{x_t^2\}_{t \in \mathbb{Z}}$, displays recognizable autocorrelations; consequently, the series $\{x_t^2\}_{t \in \mathbb{Z}}$ satisfies (1.2) and will be used for the alternative hypothesis. One of the popular models for LRD is the fractionally integrated GARCH (FIGARCH) model of Baillie *et al.* (1996). We use this as the alternative model in the power of the test and described it more details in the simulation section. The first test we consider is:

$$\begin{aligned} H_0^{(0)} &: \text{The observed data } \{x_t\}_{t \in \mathbb{Z}} \text{ follow the VS-0 model,} \\ H_1 &: \text{not } H_0. \end{aligned}$$

In order to conduct the test, consider the CUSUM test statistic based on $\{x_t\}_{t \in \mathbb{Z}}$:

$$T_n = \frac{1}{\sqrt{n}s_n} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k x_t^2 - \frac{k}{n} \sum_{t=1}^n x_t^2 \right|,$$

where n is the sample size, and s_n^2 is the estimator of the long-run variance $\sigma^2 = \sum_{h=-\infty}^{\infty} \text{Cov}(x_t^2, x_{t+h}^2)$. Note that the CUSUM test for GARCH models suffers from size distortions and low powers under finite sample size due to the strong correlation between observations. Therefore, Lee *et al.* (2015) consider the CUSUM test based on the i.i.d. innovation series $\{\epsilon_t\}$ in equation (1.1), which are uncorrelated observations. That is,

$$\hat{T}_n = \frac{1}{\sqrt{n\tau}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k \epsilon_t^2 - \frac{k}{n} \sum_{t=1}^n \epsilon_t^2 \right|,$$

where $\tau = \text{Var}(\epsilon_t^2)$. Since $\{\epsilon_t\}$ is unobservable, one can estimate the innovation series by

$$\tilde{\epsilon}_t := \frac{x_t}{\tilde{\sigma}_t(\hat{\theta})}, \quad t = 1, \dots, n, \tag{2.1}$$

where $\hat{\theta} = (\hat{\omega}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_q, \hat{\beta}_1, \dots, \hat{\beta}_p)'$ is the estimated parameter vector of GARCH models, and $\tilde{\sigma}_t^2(\theta)$ is calculated recursively from

$$\tilde{\sigma}_t^2(\theta) = \omega + \sum_{i=1}^q \alpha_i x_{t-i}^2 + \sum_{j=1}^p \beta_j \tilde{\sigma}_{t-j}^2(\theta),$$

where $\theta = (\omega, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$ is the parameter vector. Note that calculating $\tilde{\sigma}_t^2$ recursively requires the initial values for x_0^2, \dots, x_{1-q}^2 and $\tilde{\sigma}_0^2, \dots, \tilde{\sigma}_{1-p}^2$. The fGarch package in R provides the initial values to be the sample average of x_1^2, \dots, x_n^2 . Finally, the residual-based CUSUM statistic is defined as

$$\tilde{T}_n = \frac{1}{\sqrt{n\hat{\tau}}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k \tilde{\epsilon}_t^2 - \frac{k}{n} \sum_{t=1}^n \tilde{\epsilon}_t^2 \right|, \tag{2.2}$$

where

$$\hat{\tau}^2 = \frac{1}{n} \sum_{t=1}^n \tilde{\epsilon}_t^4 - \left(\frac{1}{n} \sum_{t=1}^n \tilde{\epsilon}_t^2 \right)^2$$

is a method of moment estimator of $\text{Var}(\epsilon_1^2)$.

Lee *et al.* (2015) also proved that under certain assumptions and conditions, and under $H_0^{(0)}$, as $n \rightarrow \infty$,

$$\tilde{T}_n \xrightarrow{d} \sup_{0 \leq u \leq 1} |\mathcal{B}^\circ(u)|,$$

where $\mathcal{B}^\circ(u)$ is a standard Brownian bridge. We will use this as the asymptotic distribution of the residual-based CUSUM test statistic \tilde{T}_n . According to Resnick (1992), for a standard Brownian motion, the formula for the distribution of \tilde{T}_n can be derived as

$$\Pr[\tilde{T}_n \leq v] = 1 + 2 \sum_{k=1}^{\infty} (-1)^k e^{-2k^2 v^2}, \quad v > 0 \tag{2.3}$$

and the p -values will be calculated based on (2.3) in the simulation section.

Consider a test on volatility shift for *known* R number of times with *unknown* locations against LRD after the null hypothesis of no volatility shift is rejected. Therefore, the hypothesis would become:

$H_0^{(R)}$: The observed data $\{x_t\}_{t \in \mathbb{Z}}$ follow the VS-R model,

H_1 : The observed data $\{x_t^2\}_{t \in \mathbb{Z}}$ follow the LRD model.

We follow Lee *et al.* (2015) and only consider the simple case when $R = 1$ in this paper, which is the single-volatility-shift case. The testing procedure can be easily extended to multiple volatility shifts.

Suppose the whole series consists of two different GARCH(p, q) models, namely $\{x_{1,t}\}_{t \in \mathbb{Z}}$ and $\{x_{2,t}\}_{t \in \mathbb{Z}}$, with the same standard i.i.d. innovations. That is,

$$x_t = \begin{cases} x_{1,t}, & \text{if } 1 \leq t \leq k^*, \\ x_{2,t}, & \text{if } k^* < t \leq n, \end{cases} \tag{2.4}$$

where k^* is an unknown change point and

$$\begin{cases} x_{1,t} = \sigma_{1,t} \epsilon_t, & \sigma_{1,t}^2 = \omega_1 + \sum_{i=1}^q \alpha_{1,i}^{\circ} x_{1,t-i}^2 + \sum_{j=1}^p \beta_{1,j}^{\circ} \sigma_{1,t-j}^2, \\ x_{2,t} = \sigma_{2,t} \epsilon_t, & \sigma_{2,t}^2 = \omega_2 + \sum_{i=1}^q \alpha_{2,i}^{\circ} x_{2,t-i}^2 + \sum_{j=1}^p \beta_{2,j}^{\circ} \sigma_{2,t-j}^2. \end{cases} \tag{2.5}$$

By Kokoszka and Leipus (2000), k^* can be estimated by

$$\hat{k} = \arg \max_{1 \leq k \leq n} \left| \frac{k(n-k)}{n^2} \left(\frac{1}{k} \sum_{t=1}^k x_t^2 - \frac{1}{n-k} \sum_{t=k+1}^n x_t^2 \right) \right|. \tag{2.6}$$

Note that in theoretical perspective, any change-point estimator of VS-1 model satisfying

$$\hat{k} - k^* = O_p(1).$$

will be sufficient. Theorem 6 of Kokoszka and Leipus (2000) verifies (2.7) for the change-point estimator in (2.6). Assume again that $\omega_m^\circ > 0$, $\alpha_{m,i}^\circ \geq 0$ for all $i = 1, \dots, q$ and $\beta_{m,j}^\circ \geq 0$ for all $j = 1, \dots, p$, for $m = 1, 2$. The test statistic is given by Lee *et al.* (2015) as

$$M_n = \max \{ \tilde{T}_{n,1}, \tilde{T}_{n,2} \}, \tag{2.7}$$

where

$$\tilde{T}_{n,1} = \frac{1}{\hat{\tau}_1} \hat{k}^{\frac{1}{2}} \max_{1 \leq k \leq \hat{k}} \left| \sum_{t=1}^k \tilde{\epsilon}_t^2 - \frac{k}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^2 \right|, \quad \tilde{T}_{n,2} = \frac{1}{\hat{\tau}_2} (n - \hat{k})^{\frac{1}{2}} \max_{\hat{k} < k \leq n} \left| \sum_{t=k+1}^k \tilde{\epsilon}_t^2 - \frac{k - \hat{k}}{n - \hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^2 \right|,$$

and

$$\hat{\tau}_1^2 = \frac{1}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^4 - \left(\frac{1}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^2 \right)^2, \quad \hat{\tau}_2^2 = \frac{1}{n - \hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^4 - \left(\frac{1}{n - \hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^2 \right)^2.$$

Let $\tilde{x}_{1,t} = x_t$, $t = 1, \dots, \hat{k}$, and $\tilde{x}_{2,t} = x_t$, $t = \hat{k} + 1, \dots, n$. We define recursively

$$\begin{aligned} \tilde{\sigma}_{1,t}^2(\theta_1) &= \omega_1 + \sum_{i=1}^q \alpha_{1,i} \tilde{x}_{1,t-i}^2 + \sum_{j=1}^p \beta_{1,j} \tilde{\sigma}_{1,t-j}^2(\theta_1), \quad t = 1, \dots, \hat{k}, \\ \tilde{\sigma}_{2,t}^2(\theta_2) &= \omega_2 + \sum_{i=1}^q \alpha_{2,i} \tilde{x}_{2,t-i}^2 + \sum_{j=1}^p \beta_{2,j} \tilde{\sigma}_{2,t-j}^2(\theta_2), \quad t = \hat{k} + 1, \dots, n \end{aligned}$$

with given fixed initial vectors $(\tilde{x}_{1,0}^2, \dots, \tilde{x}_{1,1-p}^2)$, $(\tilde{x}_{2,\hat{k}}^2, \dots, \tilde{x}_{2,\hat{k}-q+1}^2)$, $(\tilde{\sigma}_{1,0}^2, \dots, \tilde{\sigma}_{1,1-p}^2)$, and $(\tilde{\sigma}_{2,\hat{k}}^2, \dots, \tilde{\sigma}_{2,\hat{k}-p+1}^2)$ and parameter vectors $\theta_m = (\omega_m, \alpha_{m,1}, \dots, \alpha_{m,q}, \beta_{m,1}, \dots, \beta_{m,p})'$, $m = 1, 2$. By using these equations, Lee *et al.* (2015) then calculated $\{\tilde{\epsilon}_t, t = 1, \dots, n\}$ by

$$\tilde{\epsilon}_t = \begin{cases} \frac{x_t}{\tilde{\sigma}_{1,t}(\hat{\theta}_1)}, & t = 1, \dots, \hat{k}, \\ \frac{x_t}{\tilde{\sigma}_{2,t}(\hat{\theta}_2)}, & t = \hat{k} + 1, \dots, n. \end{cases} \tag{2.8}$$

The Gaussian QMLEs for θ_1 and θ_2 are given by

$$\hat{\theta}_1 = \operatorname{argmin}_{\theta_1 \in \Theta} \sum_{t=1}^{\hat{k}} \tilde{l}_t(\theta_1), \quad \hat{\theta}_2 = \operatorname{argmin}_{\theta_2 \in \Theta} \sum_{t=\hat{k}+1}^n \tilde{l}_t(\theta_2), \tag{2.9}$$

where Θ is a parameter space belongs to $[c_1, c_2]^{p+q+1}$ for some $0 < c_1 < c_2 < \infty$, and

$$\tilde{l}_i(\theta_m) = \frac{r_i^2}{\tilde{\sigma}_i^2(\theta_m)} + \log \tilde{\sigma}_i^2(\theta_m), \quad \tilde{\sigma}_i^2(\theta_m) = \tilde{\sigma}_{1,t}^2(\theta_m)I(t \leq \hat{k}) + \tilde{\sigma}_{2,t}^2(\theta_m)I(t > \hat{k}), \quad m = 1, 2.$$

Similarly, they also proved that if certain conditions hold, for the VS-1 model (2.5) under $H_0^{(1)}$, as $n \rightarrow \infty$,

$$M_n \xrightarrow{d} \max \left\{ \sup_{0 \leq u \leq 1} |\mathcal{B}_1^\circ(u)|, \sup_{0 \leq u \leq 1} |\mathcal{B}_2^\circ(u)| \right\},$$

where $\mathcal{B}_1^\circ(u)$ and $\mathcal{B}_2^\circ(u)$ are independent standard Brownian bridges. After calculating M_n , we can then use equation (2.3) to calculate the p -value of the test.

Note that we consider only the case when the single volatility shift is known *in priori*. The binary segmentation method (Bai, 1997) can be used to estimate the number of change points if the number of volatility shift is *unknown*.

3. Bootstrap Tests for Volatility Shifts against LRD

The use of the asymptotic test statistics for the finite sample might give size distortion and lower power. Efron (1979) introduced the bootstrap method to study the distributions of estimators (sampling distribution) and test statistics by resampling the finite sample data. In this paper, we compare the size and power of the tests presented in Section 2 using the asymptotic and bootstrap-based methods.

If the data $\{x_1, \dots, x_n\}$ are independent observations, a commonly used bootstrap procedure is to sample m observations, $\{x_1^*, \dots, x_m^*\}$, from the original data with replacement, and calculate the estimator or test statistic based on the sampled observations, say t^* . This procedure is repeated B times, and $\{t_1^*, \dots, t_B^*\}$ is obtained. Then the distribution of bootstrap replication can approximate the distribution of the estimator or test statistic under the null hypothesis. If one performs bootstrap for the test statistic, then the right-sided bootstrap p -value can be calculated as

$$\hat{p}^* = \frac{1}{B} \sum_{b=1}^B I(t_b^* \geq t),$$

where $I(\cdot)$ is the indicator function and t is the test statistic evaluated from the original data.

Time series data are autocorrelated, and therefore are not independent to each other. Simply resampling the original data might yield even worse results than the asymptotic estimation does. One method to deal with such a time series problem is to bootstrap the estimated i.i.d. innovation series $\{\tilde{\epsilon}_t\}$, and then generate recursively a bootstrap sample $\{x_t^*\}$. The algorithm for the bootstrap residual-based CUSUM test for no volatility shift against LRD is:

1. Obtain the estimated GARCH model parameters and calculate the innovation series $\{\tilde{\epsilon}_t\}$ according to equation (2.1) and compute the test statistic \tilde{T}_n by equation (2.2).
2. Construct the bootstrap innovation series $\{\tilde{\epsilon}_t^*\}$ with size equal to the original sample size n , by drawing independently with replacement from $\{\tilde{\epsilon}_t\}$.

3. Generate recursively a bootstrap sample $\{x_t^*\}$ using the estimated parameters and the bootstrap innovation series from Steps 1 and 2, respectively as:

$$x_t^* = \sigma_t^* \tilde{\epsilon}_t^*,$$

$$\sigma_t^{*2} = \hat{\omega} + \sum_{i=1}^q \hat{\alpha}_i x_{t-i}^{*2} + \sum_{j=1}^p \hat{\beta}_j \tilde{\sigma}_{t-j}^{*2},$$

with fixed initial values for $x_0^{*2}, \dots, x_{1-q}^{*2}$ and $\tilde{\sigma}_0^{*2}, \dots, \tilde{\sigma}_{1-p}^{*2}$.

4. Compute the bootstrap test statistic \tilde{T}_n^* from the bootstrap sample $\{x_t^*\}$ as follows:

$$\tilde{T}_n^* = \frac{1}{\sqrt{n\hat{\tau}^*}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k \tilde{\epsilon}_t^{**2} - \frac{k}{n} \sum_{t=1}^n \tilde{\epsilon}_t^{**2} \right|,$$

where $\{\tilde{\epsilon}_t^{**}\}$ is defined through the bootstrap sample $\{x_t^*\}$ by $\tilde{\epsilon}_t^{**} = x_t^*/\tilde{\sigma}_t^{**}$ for $t = 1, 2, \dots, n$. Here, $\hat{\theta}^* = (\hat{\omega}^*, \hat{\alpha}_1^*, \dots, \hat{\alpha}_q^*, \hat{\beta}_1^*, \dots, \hat{\beta}_p^*)'$ is the estimated parameter vector of GARCH models for $\{x_t^*\}$ and

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

with fixed initial values for $x_0^{*2}, \dots, x_{1-q}^{*2}$ and $\tilde{\sigma}_0^{**2}, \dots, \tilde{\sigma}_{1-p}^{**2}$.

5. Repeat Steps 2–4 B times to obtain $\tilde{T}_{n_1}^*, \dots, \tilde{T}_{n_B}^*$.
6. Estimate the bootstrap p -value by

$$\hat{p}^* = \frac{1}{B} \sum_{b=1}^B I(\tilde{T}_{n_b}^* \geq \tilde{T}_n),$$

where $I(\cdot)$ is the indicator function.

For the test statistic M_n for one volatility shift against LRD, the bootstrap procedure is given as:

1. Get the estimated change point \hat{k} by equation (2.6) and obtain the estimated parameters of GARCH models as defined by equation (2.9).
2. Calculate innovation series $\{\tilde{\epsilon}_t\}$ by equation (2.8) and compute the test statistic M_n as in equation (2.7).
3. Construct the bootstrap innovation series $\{\tilde{\epsilon}_t^*\}$ with size equal to the original sample size, by drawing independently with replacement from $\{\tilde{\epsilon}_t\}$.
4. Generate recursively a bootstrap sample $\{x_t^*\}$ by using the estimated change point \hat{k} , estimated parameters $\hat{\theta}_m = (\hat{\omega}_m, \hat{\alpha}_{m,1}, \dots, \hat{\alpha}_{m,q}, \hat{\beta}_{m,1}, \dots, \hat{\beta}_{m,p})'$ for $m = 1, 2$ of GARCH models and the bootstrap innovation series $\{\tilde{\epsilon}_t^*\}$ from Steps 1 and 3, respectively as:

$$x_t^* = \sigma_t^* \tilde{\epsilon}_t^*,$$

where

$$\sigma_t^{*2} = \begin{cases} \hat{\omega}_1 + \sum_{i=1}^q \hat{\alpha}_{1,i} x_{t-i}^{*2} + \sum_{j=1}^p \hat{\beta}_{1,j} \tilde{\sigma}_{1,t-j}^{*2}, & t = 1, \dots, \hat{k}, \\ \hat{\omega}_2 + \sum_{i=1}^q \hat{\alpha}_{2,i} x_{t-i}^{*2} + \sum_{j=1}^p \hat{\beta}_{2,j} \tilde{\sigma}_{2,t-j}^{*2}, & t = \hat{k} + 1, \dots, n \end{cases}$$

with fixed initial values for $x_0^{*2}, \dots, x_{1-q}^{*2}$ and $\tilde{\sigma}_0^{*2}, \dots, \tilde{\sigma}_{1-p}^{*2}$.

5. Compute the bootstrap test statistic M_n^* from the bootstrap sample $\{x_t^*\}$ as:

$$M_n^* = \max\{\tilde{T}_{n,1}^*, \tilde{T}_{n,2}^*\},$$

where

$$\tilde{T}_{n,1}^* = \frac{1}{\hat{\tau}_1^{*2}} \hat{k}^{\frac{1}{2}} \max_{1 \leq k \leq \hat{k}} \left| \sum_{t=1}^k \tilde{\epsilon}_t^{**2} - \frac{k}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^{**2} \right|, \quad \tilde{T}_{n,2}^* = \frac{1}{\hat{\tau}_2^{*2}} (n - \hat{k})^{\frac{1}{2}} \max_{\hat{k} < k \leq n} \left| \sum_{t=\hat{k}+1}^k \tilde{\epsilon}_t^{**2} - \frac{k - \hat{k}}{n - \hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^{**2} \right|,$$

and

$$\hat{\tau}_1^{*2} = \frac{1}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^{**4} - \left(\frac{1}{\hat{k}} \sum_{t=1}^{\hat{k}} \tilde{\epsilon}_t^{**2} \right)^2, \quad \hat{\tau}_2^{*2} = \frac{1}{n - \hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^{**4} - \left(\frac{1}{n - \hat{k}} \sum_{t=\hat{k}+1}^n \tilde{\epsilon}_t^{**2} \right)^2,$$

where $\{\tilde{\epsilon}_t^{**}\}$ is defined through the bootstrap sample $\{x_t^*\}$ by $\tilde{\epsilon}_t^{**} = x_t^* / \tilde{\sigma}_t^{**}$ for $t = 1, 2, \dots, n$. Here, $\hat{\theta}_m^* = (\hat{\omega}_m^*, \hat{\alpha}_{m,1}^*, \dots, \hat{\alpha}_{m,q}^*, \hat{\beta}_{m,1}^*, \dots, \hat{\beta}_{m,p}^*)'$ for $m = 1, 2$ is the estimated parameter vector of GARCH models, obtained by (2.9) based on $\{x_t^*\}$ and

$$\sigma_t^{*2} = \begin{cases} \hat{\omega}_1^* + \sum_{i=1}^q \hat{\alpha}_{1,i}^* x_{t-i}^{*2} + \sum_{j=1}^p \hat{\beta}_{1,j}^* \tilde{\sigma}_{1,t-j}^{*2}, & t = 1, \dots, \hat{k}, \\ \hat{\omega}_2^* + \sum_{i=1}^q \hat{\alpha}_{2,i}^* x_{t-i}^{*2} + \sum_{j=1}^p \hat{\beta}_{2,j}^* \tilde{\sigma}_{2,t-j}^{*2}, & t = \hat{k} + 1, \dots, n, \end{cases}$$

with fixed initial values for $x_0^{*2}, \dots, x_{1-q}^{*2}$ and $\tilde{\sigma}_0^{*2}, \dots, \tilde{\sigma}_{1-p}^{*2}$.

6. Repeat Steps 3–5 B times to obtain $M_{n_1}^*, \dots, M_{n_B}^*$.

7. Estimate the bootstrap p -value by

$$\hat{p}^* = \frac{1}{B} \sum_{b=1}^B I(M_{n_b}^* \geq M_n),$$

where $I(\cdot)$ is the indicator function.

4. Simulation Study

For empirical sizes of tests for volatility shifts against LRD, we consider the univariate GARCH(1, 1) models as

$$x_t = \sigma_t \epsilon_t,$$

$$\sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2,$$

where $\{\epsilon_t\}_{t \in \mathbb{Z}}$ is the innovation series following specific distributions. In the test for no volatility shifts against LRD, we consider two types of innovation distributions: i.i.d. standard normal and t -distribution with degrees of freedom 5, denoted by $t(5)$, following the setting from Lee *et al.* (2015). We also follow the 6 sets of GARCH(1, 1) parameters (ω, α, β) with $\alpha + \beta$ ranging from 0.5 to very close to 1. In the test for one volatility shift against LRD, i.i.d. standard normal is only considered as the innovation series. The GARCH(1, 1) parameters shift at the midpoint of the samples from $(\omega, \alpha, \beta) = (0.1, 0.1, 0.8)$ to 4 other sets with $\alpha + \beta$ ranging from 0.5 to very close to 1.

For the empirical powers, we consider only the FIGARCH(1, d , 1) models, proposed by Baillie *et al.* (1996), given as

$$x_t = \sigma_t \epsilon_t,$$

$$(1 - \beta L)\sigma_t^2 = \omega + \left((1 - \beta L) + (1 - \phi L)(1 - L)^d \right) x_t^2.$$

Here, L is the backshift operator. In the FIGARCH(1, d , 1) models, the power d of the fractional differencing operator, $(1 - L)^d$, is allowed to be a non-integer and can be expanded as

$$(1 - L)^d = \sum_{i=0}^{\infty} \pi_i L^i,$$

where

$$\pi_i = \prod_{0 \leq k \leq i} \frac{k - 1 - d}{k},$$

and $d \in (-1/2, 1/2)$. For $d > 0$, the process is long range dependent. In the simulation of empirical powers, we set FIGARCH(1, d , 1) parameters as $(\omega, \beta, \phi) = (0.6, 0.1, 0.2)$ and $(0.2, 0.2, 0.2)$, each with three different values of $d = 0.25, 0.35, 0.45$. All results are based on 1000 replications on sample sizes 500, 750, 1000 and 2000 with significant level 5%, and the bootstrap replication is set to be $B = 100$.

Table 1 presents the empirical sizes of the test for no volatility shift against LRD under different sample sizes and different types of innovations. It contains the asymptotic (columns under ‘‘Asymptotic’’) and bootstrap (columns under ‘‘Bootstrap’’) sizes for the test statistic \tilde{T}_n in (2.2) and the first bootstrap algorithm in Section 3, respectively. According to Table 1, the bootstrap method yields better sizes within acceptable range than the asymptotic ones, even $\alpha + \beta$ are very close to 1. This difference is more evident when the innovation follows a heavy-tailed distribution; in addition, the bootstrap size is closer to the nominal 5% significance level when the sample size is small, unlike asymptotic ones. When the sample sizes get larger, the bootstrap size gets similar to the asymptotic size when the innovation is normally distributed, but still better than the asymptotic ones when the innovation follows heavy-tailed distribution. Hence, the bootstrap method might be a good choice

Table 1: Empirical sizes of test for no volatility shifts under test statistic \tilde{T}_n

	(ω, α, β)	$N(0, 1)$		$t(5)$	
		Asymptotic	Bootstrap	Asymptotic	Bootstrap
$n = 500$	(0.1, 0.1, 0.8)	0.024	0.051	0.011	0.039
	(0.1, 0.1, 0.6)	0.022	0.061	0.018	0.058
	(0.1, 0.1, 0.4)	0.025	0.058	0.017	0.045
	(0.1, 0.2, 0.6)	0.024	0.042	0.014	0.030
	(0.3, 0.1, 0.8)	0.023	0.049	0.018	0.047
	(0.3, 0.1, 0.89)	0.010	0.043	0.015	0.044
$n = 750$	(0.1, 0.1, 0.8)	0.032	0.060	0.022	0.042
	(0.1, 0.1, 0.6)	0.026	0.048	0.009	0.038
	(0.1, 0.1, 0.4)	0.020	0.042	0.017	0.046
	(0.1, 0.2, 0.6)	0.034	0.048	0.036	0.052
	(0.3, 0.1, 0.8)	0.024	0.041	0.028	0.063
	(0.3, 0.1, 0.89)	0.016	0.046	0.021	0.052
$n = 1000$	(0.1, 0.1, 0.8)	0.034	0.052	0.014	0.036
	(0.1, 0.1, 0.6)	0.037	0.057	0.029	0.049
	(0.1, 0.1, 0.4)	0.036	0.052	0.021	0.043
	(0.1, 0.2, 0.6)	0.043	0.050	0.032	0.051
	(0.3, 0.1, 0.8)	0.034	0.052	0.029	0.044
	(0.3, 0.1, 0.89)	0.020	0.049	0.014	0.034
$n = 2000$	(0.1, 0.1, 0.8)	0.040	0.048	0.027	0.041
	(0.1, 0.1, 0.6)	0.047	0.058	0.028	0.046
	(0.1, 0.1, 0.4)	0.051	0.054	0.025	0.040
	(0.1, 0.2, 0.6)	0.052	0.054	0.023	0.034
	(0.3, 0.1, 0.8)	0.040	0.048	0.035	0.047
	(0.3, 0.1, 0.89)	0.027	0.044	0.035	0.047

Table 2: Empirical powers of test for no volatility shifts under test statistic \tilde{T}_n

	d	$n = 500$		$n = 750$	
		Asymptotic	Bootstrap	Asymptotic	Bootstrap
FIGARCH(1, d , 1) ($\omega = 0.6, \beta = 0.1$)	0.25	0.197	0.249	0.198	0.253
	0.35	0.271	0.324	0.261	0.328
	0.45	0.309	0.369	0.299	0.345
FIGARCH(1, d , 1) ($\omega = 0.2, \beta = 0.2$)	0.25	0.144	0.201	0.154	0.217
	0.35	0.203	0.273	0.189	0.256
	0.45	0.245	0.296	0.254	0.300
	d	$n = 1000$		$n = 2000$	
		Asymptotic	Bootstrap	Asymptotic	Bootstrap
FIGARCH(1, d , 1) ($\omega = 0.6, \beta = 0.1$)	0.25	0.342	0.379	0.558	0.573
	0.35	0.446	0.477	0.667	0.685
	0.45	0.489	0.528	0.713	0.719
FIGARCH(1, d , 1) ($\omega = 0.2, \beta = 0.2$)	0.25	0.260	0.302	0.444	0.474
	0.35	0.370	0.401	0.586	0.610
	0.45	0.433	0.453	0.631	0.651

FIGARCH = fractionally integrated generalized autoregressive conditionally heteroskedastic

when either data show the evidence of heavy tailed innovation or the sample size is small. Table 2 summarizes the power comparison in the bootstrap method that is more powerful than the asymptotic one for the cases with small samples; however, the power of both methods increases dramatically as the sample size increases and the advantage of the bootstrap diminishes.

Table 3 presents the empirical sizes of the test for one volatility shift against LRD. The results show that the bootstrap size is more assuring under the small sample sizes for test \tilde{M}_n . The asymptotic size gets close to the nominal significance level as the sample size increases. Table 4 shows the empirical

Table 3: Empirical sizes of test for one volatility shift under test statistic \tilde{M}_n

$(0.1, 0.1, 0.8)$ to $(\omega_2, \alpha_2, \beta_2)$	$n = 500$		$n = 750$	
	Asymptotic	Bootstrap	Asymptotic	Bootstrap
(0.1, 0.1, 0.6)	0.007	0.059	0.017	0.072
(0.1, 0.1, 0.4)	0.004	0.064	0.009	0.058
(0.1, 0.2, 0.6)	0.006	0.054	0.014	0.060
(0.3, 0.1, 0.8)	0.004	0.040	0.006	0.033
$(0.1, 0.1, 0.8)$ to $(\omega_2, \alpha_2, \beta_2)$	$n = 1000$		$n = 2000$	
	Asymptotic	Bootstrap	Asymptotic	Bootstrap
(0.1, 0.1, 0.6)	0.011	0.058	0.037	0.072
(0.1, 0.1, 0.4)	0.015	0.044	0.032	0.061
(0.1, 0.2, 0.6)	0.016	0.058	0.032	0.058
(0.3, 0.1, 0.8)	0.010	0.043	0.023	0.058

Table 4: Empirical powers of test for one volatility shift under test statistic \tilde{M}_n

d	$n = 500$		$n = 750$	
	Asymptotic	Bootstrap	Asymptotic	Bootstrap
FIGARCH(1, d , 1) ($\omega = 0.6, \beta = 0.1$)	0.25	0.065	0.149	0.105
	0.35	0.075	0.183	0.198
	0.45	0.084	0.198	0.229
FIGARCH(1, d , 1) ($\omega = 0.2, \beta = 0.2$)	0.25	0.039	0.124	0.089
	0.35	0.061	0.143	0.115
	0.45	0.068	0.171	0.151
d	$n = 1000$		$n = 2000$	
	Asymptotic	Bootstrap	Asymptotic	Bootstrap
FIGARCH(1, d , 1) ($\omega = 0.6, \beta = 0.1$)	0.25	0.188	0.270	0.410
	0.35	0.245	0.346	0.513
	0.45	0.301	0.450	0.550
FIGARCH(1, d , 1) ($\omega = 0.2, \beta = 0.2$)	0.25	0.115	0.200	0.272
	0.35	0.166	0.264	0.416
	0.45	0.218	0.313	0.457

FIGARCH = fractionally integrated generalized autoregressive conditionally heteroskedastic

power of test for one volatility shift. Overall, the power increases significantly when the sample size increases and the long range dependence of the series increases; in addition, the bootstrap method yields superior powers than asymptotic method when the sample size is small. This difference is barely observed as the sample size increases. In conclusion, bootstrap tests are shown to be more preferable for one volatility shift against LRD when sample sizes are small.

5. Conclusion

Lee *et al.* (2015) proposed the test statistics for volatility shifts against LRD based on the CUSUM statistic using residual series from the fitted GARCH models. In this paper, we extended it to the bootstrap methods and compared the empirical sizes and powers of the test statistics of Lee *et al.* (2015) and our bootstrap tests. We concluded that our bootstrap methods are more promising than the asymptotic ones for volatility shifts against LRD in small sample sizes, since empirical sizes are closer to the nominal significant level and empirical powers are higher.

References

Andreou, E. and Ghysels, E. (2002). Detecting multiple breaks in financial market volatility dynamics, *Journal of Applied Econometrics*, **17**, 579–600.
 Bai, J. (1997). Estimation of a change point in multiple regression models, *Review of Economics and*

- Statistics*, **79**, 551–563.
- Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **74**, 3–30.
- Beran, J. (1994). *Statistics for Long-Memory Processes*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability (Book 61), Chapman & Hall, New York.
- Berkes, I., Horváth, L., Kokoszka, P. and Shao, Q.-M. (2006). On discriminating between long-range dependence and changes in mean, *Annals of Statistics*, **34**, 1140–1165.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics*, **31**, 307–327.
- Ding, Z., Granger, C. W. and Engle, R. F. (1993). A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, **1**, 83–106.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *Annals of Statistics*, **7**, 1–26.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica: Journal of the Econometric Society*, **50**, 987–1007.
- Jach, A. and Kokoszka, P. (2008). Wavelet-domain test for long-range dependence in the presence of a trend, *Statistics*, **42**, 101–113.
- Kokoszka, P. and Leipus, R. (2000). Change-point estimation in ARCH models, *Bernoulli*, **6**, 513–539.
- Kuswanto, H. (2011). A new simple test against spurious long memory using temporal aggregation, *Journal of Statistical Computation and Simulation*, **81**, 1297–1311.
- Lee, S., Tokutsu, Y. and Maekawa, K. (2004). The cusum test for parameter change in regression models with ARCH errors, *Journal of the Japan Statistical Society*, **34**, 173–188.
- Lee, T., Kim, M. and Baek, C. (2015). Tests for volatility shifts in GARCH against long-range dependence, *Journal of Time Series Analysis*, **36**, 127–153.
- Qu, Z. (2011). A test against spurious long memory, *Journal of Business and Economic Statistics*, **29**, 423–438.
- Resnick, S. I. (1992). *Adventures in Stochastic Processes*, Birkhäuser, Boston.
- Zhang, A., Gabrys, R. and Kokoszka, P. (2007). Discriminating between long memory and volatility shifts, *Austrian Journal of Statistics*, **36**, 253–275.