A hybrid bootstrap approach to unit root tests

Guodong Li, Chenlei Leng and Chih-Ling Tsai

University of Hong Kong, National University of Singapore

and University of California at Davis

February 6, 2014

Abstract

This paper proposes a hybrid bootstrap approach to approximate the augmented Dickey-Fuller test by perturbing both the residual sequence and the minimand of the objective function. Since innovations can be dependent, this allows the inclusion of conditional heteroscedasticity models. The new bootstrap method is also applied to least absolute deviation-based unit root test statistics, which are efficient in handling heavy-tailed time series data. The asymptotic distributions of resulting bootstrap tests are presented, and Monte Carlo studies demonstrate the usefulness of the proposed tests.

Some keywords: Bootstrap; Brownian motion; Least absolute deviation; Unit root test

1 Introduction

In time series analysis, unit root tests have been widely studied under various scenarios, see Phillips (1987), Phillips and Perron (1988), and Fuller (1996), among others. Because the limiting distributions of unit root tests usually involve Brownian motion, it becomes difficult to calculate the critical values in practice. Hence, Dickey (1976) employed the Monte Carlo method to construct critical values for the Dickey-Fuller (1979) test. Due to the introduction of more powerful computing equipment, the bootstrap technique originally proposed by Efron (1979) has attracted increasing attention in approximating the null distributions of the Dickey-Fuller (DF) and the augmented DF (ADF) tests, see Paparoditis and Politis (2005), Palm et al. (2008), Phillips (2010) and references therein.

In the context of bootstrap unit root tests, the sieve bootstrap is one of the most popular approaches being considered. It employs an autoregressive (AR) model in order to remove the correlation structure of time series, and then re-samples the resulting residuals, see Chang and Park (2003) and Paparoditis and Politis (2005). Since the time order of residuals is destroyed by the resampling operation, it is usually assumed that the innovations are independently and identically distributed (*i.i.d*), see Paparoditis and Politis (2005). To allow weaker assumptions on the innovation structure of the process, Paparoditis and Politis (2003) proposed a residual-based block (RBB) bootstrap method for unit root tests in which the blocks of residuals are resampled. However, selecting the block size is a challenging task; see Palm et al. (2008). Recently, Cavaliere and Taylor (2009b) applied a wild bootstrap approach to unit root processes with a very general class of non-stationary heteroscedastic innovations, and Cavaliere and Taylor (2009a) proposed wild bootstrap implementations for the M unit root tests of Ng and Perron (2001). In sum, the above bootstrap methods are basically based on residuals.

In contrast to bootstrapping residuals, Jin et al. (2001) proposed an alternative bootstrap method by perturbing the minimand of the objective function, and later Chatterjee and Bose (2005) introduced an approach by perturbing the estimating equations. It is worth noting that the above two bootstrap methods only focused on parameter estimations. Monte Carlo studies in Chatterjee and Bose (2005) show that their technique is superior to the residual bootstrap and the wild bootstrap for three models: heteroscedastic time series, generalized linear model, and nonlinear regression. Recently, Chen et al. (2008) demonstrated that it also works well for testing the linear hypothesis. These findings motivate us to apply this new approach for the ordinary least squares (OLS) based unit root tests.

In practice, many financial and economic time series are heavy-tailed, and the least absolute deviation (LAD) approach is usually used to deal with these types of data (e.g., Peng and Yao, 2003; Li and Li, 2008). Herce (1996) studied the LAD-based unit root tests, and Moreno and Romo (2000) provided a bootstrapping approximation to the null distribution for *i.i.d.* innovations. In addition, Li and Li (2009) discussed the LAD estimation for the unit root process with GARCH innovations. Unlike the case of the OLS, the asymptotic distributions of the estimated unit roots have a very complicated form, and some strong conditions such as symmetry are needed, see Li and Li (2009). This also inspires us to propose a novel bootstrap method for the LAD-based unit root test, which does not require those strong conditions, see Remark 5 at section 3 for details.

The aim of this paper is to propose a hybrid bootstrap (HB) approach for unit root tests. Specifically, we combine the perturbation of residuals, as in the wild bootstrap, with the perturbation of the minimand of the objective function, as in Jin et al. (2001), to construct easily implemented bootstrap unit root tests for time series with uncorrelated but possibly dependent innovations. Accordingly, the HB method is applicable for the time-varying conditional variance (Engle, 1982; Bollerslev, 1986), which is an important feature in financial time series and has been well discussed in unit root tests, see Seo (1999), Chang and Park (2002) and Ling and Li (2003). It is noteworthy that the bootstrap method in Jin et al. (2001) itself will not meet the intended purpose by providing the approximating distribution of the normality instead of the desired functional of Brownian motion, see Remark 2 at section 2 for details.

The rest of this paper is organized as follows. Section 2 introduces the hybrid bootstrap unit root test via OLS estimators, while section 3 develops the hybrid bootstrap unit root test via LAD estimators. Theoretical properties of resulting HB tests are also obtained. Subsequently, Monte Carlo studies are presented in section 4, and section 5 gives a final conclusion. In this paper, all detailed proofs are relegated to the Appendix. In addition, $\|\cdot\|$ denotes the Euclidean norm of a vector or a matrix, $o_p(1)$ denotes a series of random variables (vectors) converging to zero in probability, $O_p(1)$ denotes a series of random variables (vectors) that are bounded in probability, $D = D[0,1]$ denotes the space of functions on $[0,1]$, which is defined and equipped with the Shorokhod topology (Billingsley, 1999), and $\Rightarrow$ denotes weak convergence on D.

2 A hybrid bootstrap unit root test via OLS estimators

Consider the following process,

$$\Delta y_t = \phi y_{t-1} + u_t, \quad u_t = \pi(L)e_t, \tag{1}$$

where $\Delta y_t = y_t - y_{t-1}$, L is a back-shift operator, $\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j$, and the innovations $\{e_t\}$ are uncorrelated with mean zero and unconditional variance σ^2 for $t = 1, \cdots, n$. To study the theoretical properties of the tests, we assume that the innovation sequence satisfies the assumptions given below.

Assumption 1. *The polynomial $\pi(z) \neq 0$ for all $|z| \leq 1$, $\sum_{j=0}^{\infty} j|\pi_j| < \infty$, the sequence $\{e_t\}$ is strictly stationary and ergodic with $E(e_t^4) < \infty$.*

Under the above assumption, $\{u_t\}$ is a stationary and invertible general linear process. In addition, an important special case of $\{e_t\}$ is that of conditionally heteroscedastic innovations, such as the ARCH-type processes (Engle, 1982; Bollerslev, 1986). Moreover, $\phi = 0$ in (1) corresponds to the presence of a unit root, while that of $-2 < \phi < 0$ leads to the stationarity of $\{y_t\}$. Accordingly, given observations $y_1, ..., y_n$ from model (1) with initial value $y_0 = 0$, we consider the following unit root test,

$$H_0 : \phi = 0 \quad \text{vs} \quad H_1 : -2 < \phi < 0.$$

To implement the ADF test, an AR structure is employed to approximate the first order dependence of $\{u_t\}$. As a result, we consider the auxiliary AR model as follows,

$$\Delta y_t = \phi y_{t-1} + \sum_{i=1}^{p} \psi_i \Delta y_{t-i} + e_{t,p}, \tag{2}$$

where p is a function of n, and $e_{t,p}$ depends on p. Adopting Said and Dickey's (1984) approach, we further assume that the order p satisfies the following assumption.

Assumption 2. *The order p is such that $p \to \infty$ and $n^{-1/3}p \to 0$ as $n \to \infty$.*

It is noteworthy that, under Assumption 1, the stochastic process $\{u_t\}$ has the AR representation of $\psi(L)u_t = e_t$, where $\psi(z) = 1 - \sum_{j=1}^{\infty} \psi_j z^j$ and $\sum_{j=1}^{\infty} j|\psi_j| < \infty$, see Chang and Park (2002). Furthermore, under the null hypothesis of $\phi = 0$, we have $\Delta y_t = u_t$ and $u_t = \sum_{i=1}^{p} \psi_i u_{t-i} + e_{t,p}$, where $e_{t,p} = e_t + \sum_{i=p+1}^{\infty} \psi_i u_{t-i}$.

To construct the test statistic, we obtain OLS estimators given below by fitting the data with model (2).

$$(\widehat{\phi}_n, \widehat{\psi}_1, ..., \widehat{\psi}_p)' = \text{argmin} \sum_{t=p+2}^{n} (\Delta y_t - \phi y_{t-1} - \sum_{i=1}^{p} \psi_i \Delta y_{t-i})^2.$$

Then, the ADF test statistic is as follows,

$$S_n = n\widehat{\phi}_n/(1 - \sum_{i=1}^{p} \widehat{\psi}_i).$$

Under Assumptions 1 and 2, it can be shown that $n\widehat{\phi}_n = O_p(1)$, $\|\widehat{\Psi} - \Psi\| = o_p(n^{-1/6})$, and

$$S_n \Rightarrow \frac{\int_0^1 B(\tau)dB(\tau)}{\int_0^1 B^2(\tau)d\tau}, \tag{3}$$

where $\Psi = (\psi_1, ..., \psi_p)'$, $\widehat{\Psi} = (\widehat{\psi}_1, ..., \widehat{\psi}_p)'$, and $B(t)$ is a standard Brownian motion process, see Chang and Park (2002). The asymptotic distribution in (3) is a function of Brownian motion, and Chang and Park (2003) suggested a sieve bootstrap to approximate it with $\{e_t\}$ being *i.i.d.* random variables.

For the conditionally heteroscedastic innovations, however, the sieve bootstrap by using the resampled residuals via the conditionally *i.i.d.* innovations assumption may fail to approximate the quantities $\widehat{\phi}_n$ and $\widehat{\Psi}$ in the test statistic S_n, see Goncalves and Kilian (2007). This motivates us to propose a hybrid bootstrap method for unit root tests via perturbing both the residual sequence and the minimand of the objective function to approximate the asymptotic distribution in (3). After fitting model (2), denote the residual sequence by $\{\widehat{e}_{t,p}, 1 \le t \le n\}$, where $\widehat{e}_{t,p} = 0$ for $1 \le t \le p+1$. We then employ the wild bootstrap approach to perturb the residuals by $\{\omega_t\}$, a sequence of *i.i.d.* non-negative random variables with mean one, variance one, and $E(\omega_t^4) < \infty$. It results in a new residual sequence $\{e_t^*\}$ with $e_t^* = (\omega_t - 1)\widehat{e}_{t,p}$ for $1 \le t \le n$. Let

$$y_t^* = y_{t-1}^* + \sum_{i=1}^{p} \widehat{\psi}_i \Delta y_{t-i}^* + e_t^*,$$

where $1 \leq t \leq n$, $\Delta y_t^* = y_t^* - y_{t-1}^*$ and the initial values of $y_1^*, ..., y_{p+1}^*$ can be set to zero. By Theorem 18.2 of Billingsley (1999) and the Beveridge-Nelson representation, we can show that, conditional on $y_1, ..., y_n$,

$$\frac{1}{\sigma\sqrt{n}} y_{[n\tau]}^* = (1 - \sum_{i=1}^{p} \widehat{\psi}_i)^{-1} \cdot \frac{1}{\sigma\sqrt{n}} \sum_{t=1}^{[n\tau]} e_t^* + R_n \Rightarrow B^*(\tau),$$

in probability, where $0 \leq \tau \leq 1$, $[n\tau]$ is the integral part of $n\tau$, $E\{|R_n||y_1, ..., y_n\} = O_p(n^{-1/2})$ and $B^*(\tau)$ is a standard Brownian motion process. From the proof of Theorem 1, the constructed sequence $\{y_t^*\}$ is only involved in the above asymptotic distribution. Hence, we may alternatively generate it by $y_t^* = (1 - \sum_{i=1}^{p} \widehat{\psi}_i)^{-1} \sum_{i=1}^{t} e_i^*$, and the asymptotic result in Theorem 1 still holds.

We next follow Jin et al.'s (2001) approach and obtain two auxiliary estimators by minimizing their corresponding objective functions,

$$(\widehat{\phi}_{1n}^*, \widehat{\Psi}_{1n}^{*\prime})' = \text{argmin} \sum_{t=p+2}^{n} (\Delta y_t - \phi y_{t-1}^* - \sum_{i=1}^{p} \psi_i \Delta y_{t-i} - \widehat{\phi}_n y_{t-1})^2 \tag{4}$$

and

$$(\widehat{\phi}_{2n}^*, \widehat{\Psi}_{2n}^{*\prime})' = \text{argmin} \sum_{t=p+2}^{n} \omega_t (\Delta y_t - \phi y_{t-1}^* - \sum_{i=1}^{p} \psi_i \Delta y_{t-i} - \widehat{\phi}_n y_{t-1})^2. \tag{5}$$

Note that ω_t's are all nonnegative, and then equation (5) can be treated as a weighted OLS estimation with random weights. The hybrid bootstrap approach yields the quantity $S_n^* = n(\widehat{\phi}_{2n}^* - \widehat{\phi}_{1n}^*)/(1 - \sum_{i=1}^{p} \widehat{\psi}_i)$; its theoretical property is given below.

Theorem 1. *Under H_0 or H_1, if Assumptions 1 and 2 hold, then, conditional on $y_1, ..., y_n$,*

$$S_n^* = \frac{n(\widehat{\phi}_{2n}^* - \widehat{\phi}_{1n}^*)}{1 - \sum_{i=1}^{p} \widehat{\psi}_i} \Rightarrow \frac{\int_0^1 B^*(\tau) dB^*(\tau)}{\int_0^1 B^{*2}(\tau) d\tau},$$

in probability, where $B^*(t)$ is a standard Brownian motion process.

The above theorem together with equation (3) allows us to approximate the null distribution of the test statistic S_n by generating B bootstrap samples of *i.i.d.* nonnegative random variables $\{\omega_t\}$ with mean one and variance one. The detailed procedure of bootstrapping unit root tests is given as follows:

(a) Calculate the value of $S_n = n\widehat{\phi}_n/(1 - \sum_{i=1}^{p} \widehat{\psi}_i)$ by fitting $\{y_t, t = 1, ..., n\}$ with model (2);

(b) Generate an $i.i.d.$ sequence $\{\omega_t, t = 1, ..., n\}$, and then calculate the value of $S^*_{n(1)} = n(\widehat{\phi}^*_{2n} - \widehat{\phi}^*_{1n})/(1 - \sum_{i=1}^{p} \widehat{\psi}_i)$;

(c) Repeat step (b), and obtain $S^*_{n(2)}, ..., S^*_{n(B)}$;

(d) Compute the empirical α-percentiles of $\{S^*_{n(i)}, i = 1, ..., B\}$, denoted by $S_B^{*\alpha}$, and reject the null hypothesis if $S_n < S_B^{*\alpha}$, where α is the predetermined significance level for a one-side test.

To select the order p at equation (2) in practice, we may consider the modified Akaike information criterion (MAIC) in Ng and Perron (2001),

$$MAIC(p) = \log(\widehat{\sigma}_p^2) + 2(p + 1 + \tau_p)/(n - p_{\max} - 1), \tag{6}$$

where $0 \leq p \leq p_{\max}$, $\widehat{\sigma}_p^2 = (n - p_{\max} - 1)^{-1} \sum_{t=p_{\max}+2}^{n} \widehat{e}_{p,t}^2$, $\widehat{e}_{p,t} = \Delta y_t - \widehat{\phi}_n y_{t-1} - \sum_{i=1}^{p} \widehat{\psi}_i \Delta y_{t-i}$, and $\tau_p = \widehat{\sigma}_p^{-2} \widehat{\phi}_n^2 \sum_{t=p_{\max}+2}^{n} y_{t-1}^2$. As in Ng and Perron (2001) and Cavaliere and Taylor (2009b), the maximum lag p_{max} can be set to $[12(n/100)^{1/4}]$, where $[x]$ is the integer part of x.

Remark 1. There are two most common types of bootstrapping unit root tests in the literature: residual-based and difference-based tests, see Paparoditis and Politis (2005) and Palm et al. (2008). Strictly speaking, S_n^* is neither of them. Since S_n^* involves the residuals and parameter estimator by fitting model (2) via the OLS approach, we can view it as a residual-based test. Alternatively, we can construct the bootstrapping test by removing $\widehat{\phi}_n y_{t-1}$ from (4) and (5), and the same asymptotic distribution is expected under H_0. However, the time series $\{\Delta y_t\}$ is not invertible under H_1. Accordingly, it may seriously deteriorate the power of test, as described for the difference-based tests in Paparoditis and Politis (2005).

Remark 2. It seems natural to employ Jin et al.'s (2001) approach to directly approximate the asymptotic distribution of S_n in equation (3). For the sake of illustration, consider that $\{u_t\}$ in model (1) are $i.i.d.$ random variables with mean zero and variance

σ^2, and assume that $p = 0$. Then, under H_0, we have

$$S_n = n\widehat{\phi}_n \Rightarrow \frac{\int_0^1 B(\tau)dB(\tau)}{\int_0^1 B^2(\tau)d\tau}.$$

Following Jin et al.'s (2001) approach, we use the quantity $n(\widehat{\phi}_n^* - \widehat{\phi}_n)$ to approximate the distribution of S_n, where

$$\widehat{\phi}_n^* = \operatorname{argmin} \sum_{t=2}^{n} \omega_t (\Delta y_t - \phi y_{t-1})^2.$$

However, under H_0 and by the proof of Theorem 1, we can show that, conditional on $y_1, ..., y_n$,

$$n(\widehat{\phi}_n^* - \widehat{\phi}_n) = \frac{n^{-1}\sum_{t=2}^{n} \omega_t u_t y_{t-1}}{n^{-2}\sum_{t=2}^{n} \omega_t y_{t-1}^2} - \frac{n^{-1}\sum_{t=2}^{n} u_t y_{t-1}}{n^{-2}\sum_{t=2}^{n} y_{t-1}^2} \Rightarrow N(0, \sigma^2)$$

in probability. As a result, this direct approach does not meet our intended purpose, which motivates us to propose the hybrid bootstrap approach.

To make the hybrid bootstrap approach more practical, we consider a trend function in the model, i.e. the observed time series $\{z_t\}$ is generated by $z_t = \mu_t'\beta + y_t$, where $\{y_t\}$ is defined as in (1), and $\mu_t = 1$ for the constant trend and $\mu_t = (1, t)'$ for the linear trend. As in Elliott et al. (1996), Ng and Perron (2001) and Cavaliere and Taylor (2009a), we employ the local generalized least squares (GLS) method to de-trend the data, i.e. $\widehat{z}_t = z_t - \mu_t'\widehat{\beta}_{GLS}$, where $\widehat{\beta}_{GLS}$ is the OLS estimator for the regression of $\widetilde{z}_t = z_t - (1-\bar{c}/n)z_{t-1}$ on $\widetilde{\mu}_t = \mu_t - (1-\bar{c}/n)\mu_{t-1}$ with $z_0 = 0$. For the 5% significance level, the value of $\bar{c}$ can be set to 7.0 for the constant trend, and 13.5 for the linear trend, see Cavaliere and Taylor (2009a). We then can calculate the ADF test statistic $S_n = n\widehat{\phi}_n/(1 - \sum_{i=1}^{p} \widehat{\psi}_i)$ by replacing $\{y_t\}$ at (2) with $\{\widehat{z}_t\}$. To approximate the distribution of S_n, we first de-trend the bootstrapped sample $\{y_t^*\}$ via the local GLS method, and denote the resulting residuals by $\{\widehat{z}_t^*\}$. The values of $\widehat{\phi}_{1n}^*$ and $\widehat{\phi}_{2n}^*$ can be obtained from the two auxiliary estimations at (4) and (5) with $\{y_t\}$ and $\{y_t^*\}$ replaced respectively by $\{\widehat{z}_t\}$ and $\{\widehat{z}_t^*\}$. Let $S_n^* = n(\widehat{\phi}_{2n}^* - \widehat{\phi}_{1n}^*)/(1 - \sum_{i=1}^{p} \widehat{\psi}_i)$, and the mathematical justification is given as follows.

Corollary 1. *Suppose that Assumptions 1 and 2 are satisfied. If H_0 holds, then $S_n \Rightarrow \int_0^1 B_C(\tau)dB_C(\tau)/\int_0^1 B_C^2(\tau)d\tau$ for the constant trend, and $S_n \Rightarrow \int_0^1 B_L(\tau)dB_L(\tau)/\int_0^1 B_L^2(\tau)d\tau$ for the linear trend, where $B_C(\tau) = B(\tau) - \int_0^1 B(\tau)d\tau - \bar{c}^{-1}B(1)$, $B_L(\tau) = B(\tau) -$*

$\nu_1(B(\tau),\bar{c}) - \tau\nu_2(B(\tau),\bar{c})$, $B(\tau)$ *is a standard Brownian motion process,* $\bar{c}$ *is a constant,*

$$\nu_1(B(\tau),\bar{c}) = \frac{6+4\bar{c}}{\bar{c}^2}\left[B(1) + \bar{c}\int_0^1 B(\tau)d\tau\right] - \frac{12+6\bar{c}}{\bar{c}^2}\left[\int_0^1 \tau dB(\tau) + \bar{c}\int_0^1 \tau B(\tau)d\tau\right],$$

and

$$\nu_2(B(\tau),\bar{c}) = -\frac{6}{\bar{c}}\left[B(1) + \bar{c}\int_0^1 B(\tau)d\tau\right] + \frac{12}{\bar{c}}\left[\int_0^1 \tau dB(\tau) + \bar{c}\int_0^1 \tau B(\tau)d\tau\right].$$

Corollary 2. *Suppose that Assumptions 1 and 2 are satisfied. If* H_0 *or* H_1 *holds, then, conditional on* $y_1, ..., y_n$, $S_n^* \Rightarrow \int_0^1 B_C^*(\tau)dB_C^*(\tau)/\int_0^1 B_C^{*2}(\tau)d\tau$ *in probability for the constant trend, and* $S_n^* \Rightarrow \int_0^1 B_L^*(\tau)dB_L^*(\tau)/\int_0^1 B_L^{*2}(\tau)d\tau$ *in probability for the linear trend, where* $B_C^*(\tau) = B^*(\tau) - \int_0^1 B^*(\tau)d\tau - \bar{c}^{-1}B^*(1)$, $B_L^*(\tau) = B^*(\tau) - \nu_1(B^*(\tau),\bar{c}) - \tau\nu_2(B^*(\tau),\bar{c})$, *and* $B^*(t)$ *is a standard Brownian motion process.*

The proofs of the above two corollaries are similar to those of Theorem 3.6 in Chang and Park (2002) and Theorem 1 in this section, respectively, and we give their details in a separated supplementary file.

3 A hybrid bootstrap unit root test via LAD estimators

In time series analysis, it is not unusual to encounter heavy-tailed observations. Accordingly, the OLS estimators can be sensitive to outliers and the resulting test statistics may not be accurate and powerful. This motivates us to extend the hybrid bootstrap approach from the previous section to LAD-based unit root tests. To this end, we now consider the AR unit root process,

$$\Delta y_t = \phi y_{t-1} + \sum_{i=1}^{p} \psi_i \Delta y_{t-i} + e_t, \tag{7}$$

where p is a known non-negative integer, $e_t = \sigma_t\varepsilon_t$ for $1 \le t \le n$, $\{\varepsilon_t\}$ are *i.i.d.* random variables with mean zero and variance one, and $\sigma_t > 0$ is measurable with respect to the information set $\sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, ...)$. Note that model (7) is a special case of model (1). To investigate theoretical properties of LAD-based unit root tests, we introduce the following assumption.

Assumption 3. *The sequence* $\{\sigma_t^2\}$ *is strictly stationary and ergodic with* $E(\sigma_t^2) < \infty$. *The median of* ε_t *is equal to zero, the density function* $f(x)$ *of* ε_t *is continuous at the origin, and* $E(\varepsilon_t^2) < \infty$.

Remark 4. The unit root process (7) requires the mean of ε_t to be zero, and the above assumption further assumes that its median is zero. These conditions restrict the asymmetry of ε_t to some extent (Engle and Gonzalez-Rivera, 1991). It is noteworthy that the LAD approach attempts to estimate the conditional median, and the term $m_\varepsilon \sigma_t$ is involved in the structure of the conditional median when the quantity m_ε, the median of ε_t, is not zero. For example, if $\sigma_t^2 = 0.5 + 0.6\Delta y_{t-1}^2$, then the conditional median of Δy_t is

$$\text{median}(\Delta y_t) = \phi y_{t-1} + \sum_{i=1}^{p} \psi_i \Delta y_{t-i} + m_\varepsilon \sqrt{0.5 + 0.6\Delta y_{t-1}^2}.$$

Hence, the restriction here is necessary for a general form of the conditional variance σ_t^2. By contrast, if we assume that $\sigma_t = c$ almost surely, for a constant c (i.e., there exists only a constant $c \cdot m_\varepsilon$ involved in the structure of the conditional median), then we can relax the restriction of both mean and median to zero. In this case, however, the innovations $\{e_t\}$ becomes *i.i.d.* so that the conditional heteroscedasticity is excluded from the model setting.

For model (7), the hypotheses of the unit root test are

$$H_0 : \phi = 0 \quad \text{vs} \quad H_1 : \phi_{\min} < \phi < 0,$$

where $\phi_{\min}$ is the inferior limit of ϕ such that model (7) is stationary, see Paparoditis and Politis (2005). Let $\theta = (\phi, \psi_1, ..., \psi_p)'$, and then we obtain the LAD estimator of θ as follows,

$$\widetilde{\theta}_n = (\widetilde{\phi}_n, \widetilde{\psi}_1, ..., \widetilde{\psi}_p)' = \text{argmin} \sum_{t=p+2}^{n} |\Delta y_t - \phi y_{t-1} - \sum_{i=1}^{p} \psi_i \Delta y_{t-i}|. \tag{8}$$

Accordingly, the LAD-based ADF test statistic is

$$L_n = n\widetilde{\phi}_n / (1 - \sum_{i=1}^{p} \widetilde{\psi}_i).$$

Furthermore, let $x_t = (\Delta y_{t-1}, ..., \Delta y_{t-p})'$, $\Sigma_0 = E(\sigma_t^{-1})$, $\Sigma_1 = E(\sigma_t^{-1} x_t)$, $\Sigma_2 = E(\sigma_t^{-1} x_t x_t')$, and

$$\Omega = \begin{pmatrix} E(e_t)^2 & E(|e_t|) & E(|e_t|x_t') \\ E(|e_t|) & 1 & E(x_t') \\ E(|e_t|x_t') & E(x_t') & E[x_t x_t'] \end{pmatrix}.$$

Then, we obtain the asymptotic distribution of L_n given below.

Theorem 2. *Under H_0, if Assumption 3 holds, then*

$$L_n \Rightarrow \frac{1}{2f(0)} \cdot \frac{\int W_1(\tau)dW_2(\tau) - \Sigma_1'\Sigma_2^{-1}W_3(1)\int W_1(\tau)d\tau}{\Sigma_0 \int W_1^2(\tau)d\tau - \Sigma_1'\Sigma_2^{-1}\Sigma_1(\int W_1(\tau)d\tau)^2},$$

where $\boldsymbol{W}(\tau) = [W_1(\tau), W_2(\tau), W_3'(\tau)]'$ is a $(p+2)$-dimensional Brownian motion process with covariance matrix $\tau\Omega$.

Remark 5. Although the above theorem yields a similar result to equation (3) in Li and Li (2009), the structure of conditional variance σ_t^2 is not required here. In practice, the computation of asymptotic distribution in Theorem 2 is very complicated, and the symmetry of ε_t is usually assumed (e.g., see Li and Li 2009). Under the symmetry condition, $\Sigma_1 = 0$ and

$$L_n \Rightarrow \frac{1}{2f(0)\Sigma_0} \cdot \frac{\int W_1(\tau)dW_2(\tau)}{\int W_1^2(\tau)d\tau}.$$

However, it is known that the asymmetry and the heavy tails are two important features in financial time series, see Engle and Gonzalez-Rivera (1991). Furthermore, the density of ε_t, $f(\cdot)$, is involved in the asymptotic distribution of L_n, and it is difficult to provide a consistent estimator for the quantity $f(0)$ without assuming a parametric structure for the conditional variance σ_t^2. The above considerations motivate us to employ the hybrid bootstrap approach to approximate the asymptotic distribution in Theorem 2.

Let $\{\widetilde{e}_t, 1 \le t \le n\}$ be the residual sequence from model (7) by the LAD approach, and $y_t^* = (1 - \sum_{i=1}^p \widetilde{\psi}_i)^{-1} \sum_{i=1}^t (\omega_i - 1)\widetilde{e}_i$. Employing the same hybrid bootstrap approach as that in section 2, we obtain

$$L_n^* = \frac{n(\widetilde{\phi}_{2n}^* - \widetilde{\phi}_{1n}^*)}{1 - \sum_{i=1}^p \widetilde{\psi}_i},$$

where

$$\widetilde{\theta}^*_{1n} = (\widetilde{\phi}^*_{1n}, \widetilde{\Psi}^{*\prime}_{1n})' = \text{argmin} \sum_{t=p+2}^{n} |\Delta y_t - \phi y^*_{t-1} - \sum_{i=1}^{p} \psi_i \Delta y_{t-i} - \widetilde{\phi}_n y_{t-1}|,$$

$$\widetilde{\theta}^*_{2n} = (\widetilde{\phi}^*_{2n}, \widetilde{\Psi}^{*\prime}_{2n})' = \text{argmin} \sum_{t=p+2}^{n} \omega_t |\Delta y_t - \phi y^*_{t-1} - \sum_{i=1}^{p} \psi_i \Delta y_{t-i} - \widetilde{\phi}_n y_{t-1}|,$$

and $\widetilde{\phi}_n$ is the LAD estimator from (8). We next obtain the theoretical property of L^*_n.

Theorem 3. *Under H_0 or H_1, if Assumption 3 holds, then, conditional on $y_1, ..., y_n$,*

$$L^*_n \Rightarrow \frac{1}{2f(0)} \cdot \frac{\int W^*_1(\tau) dW^*_2(\tau) - \Sigma'_1 \Sigma_2^{-1} W^*_3(1) \int W^*_1(\tau) d\tau}{\Sigma_0 \int W^{*2}_1(\tau) d\tau - \Sigma'_1 \Sigma_2^{-1} \Sigma_1 (\int W^*_1(\tau) d\tau)^2}$$

in probability, where $\boldsymbol{W}^*(\tau) = [W^*_1(\tau), W^*_2(\tau), W^{*\prime}_3(\tau)]'$ is a $(p+2)$-dimensional Brownian motion process with covariance matrix $\tau\Omega$.

The asymptotic distribution in the above theorem is the same as that in Theorem 2, although $\mathbf{W}^*(\tau)$ and $\mathbf{W}(\tau)$ are two different Brownian motion processes. Hence, Theorems 2 and 3 allow us to apply a bootstrap procedure similar to that in Section 2 to obtain the LAD-based bootstrap unit root test via L^*_n. Accordingly, we do not need to calculate the quantities Σ_0, Σ_1, Σ_2, and $f(0)$ in Theorem 2, which mitigates the complicated computation. To select the order p in model (7), we adapt MAIC at (6) for the LAD approach by replacing $\widehat{\sigma}_p$ and τ_p respectively by $\tilde{\sigma}_p$ and $\tilde{\tau}_p$, where $\tilde{\sigma}_p = (n - p_{\max} - 1)^{-1} \sum_{t=p_{\max}+2}^{n} |\tilde{e}_{p,t}|$ and $\tilde{\tau}_p = \tilde{\sigma}_p^{-2} (\sum_{t=p_{\max}+2}^{n} |\tilde{\phi}_n y_{t-1}|)^2$.

4 Simulation studies

We conduct two Monte Carlo experiments. The first one aims to evaluate the finite sample performance of the proposed bootstrap approach and the second one aims to make comparisons with five other bootstrapping unit root tests. In both experiments, the sample size is set to $n = 100$, 200 or 300, and the three commonly used significance levels, 1%, 5% and 10%, are employed. The number of replications is fixed at 1000, and the number of bootstrapped samples is $B = 1000$.

4.1 Finite sample performance of HB tests

We now conduct Monte Carlo experiments to evaluate the finite sample performance of the proposed tests, S_n^* and L_n^*. The generating process is given as follows,

$$\Delta y_t = \phi y_{t-1} + e_t, \tag{9}$$

$$e_t = \varepsilon_t h_t^{1/2}, \quad \text{and} \quad h_t = 0.1 + 0.2e_{t-1}^2 + 0.7h_{t-1}, \tag{10}$$

where $\{\varepsilon_t\}$ are *i.i.d.* standard normal random variables. We consider four distributions for the perturbing sequence $\{\omega_t\}$: (i) the standard exponential distribution, (ii) the Rademacher distribution, which takes the value 0 or 2, each with probability 0.5, see Li and Li (2011), (iii) Mammen's two-point distribution, which takes the value $(-\sqrt{5}+3)/2$ with probability $(\sqrt{5}+1)/2\sqrt{5}$ and the value $(\sqrt{5}+3)/2$ with probability $1-(\sqrt{5}+1)/2\sqrt{5}$, see Mammen (1993), and (iv) a mixture of the distributions in (i) and (ii) with mixing probability 0.5. The third-order central moments of the distributions in (iii) and (iv) are equal to one, which may provide a better limiting distribution in Section 3 (e.g., see Liu 1988). For the sake of simplicity, we set the order p in equations (2) and (7) to be zero.

Table 1 presents the rejection rates of the test S_n^*. Under the null hypothesis with $\phi = 0.0$, the rejection rates are all close to the corresponding nominal levels across all four perturbation distributions, even in the small sample size of $n = 100$. Under the alternative hypothesis with $\phi < 0.0$, these four perturbation distributions provide comparable empirical powers, see Mammen (1993) for similar findings. It is not surprising that the power becomes larger as the sample size increases or ϕ gets smaller. Since the LAD-based test L_n^* yields similar results, we omitted them. It is worth mentioning that S_n^* is generally superior to L_n^*. To make further comparisons, we follow the same model structure, using (9) and (10), to generate sample data. Since the four perturbing distributions show similar results, we only consider the Mammen's two-point distribution for the perturbing sequence $\{\omega_t\}$. In addition, the innovations $\{\varepsilon_t\}$ are *i.i.d.* Student's $t(3)$ random variables, which have been standardized to have mean zero and variance one. Table 2 shows that S_n^* is inferior to L_n^* for heavy-tailed innovations, GARCH-$t(3)$. This suggests that one could consider the LAD-based test rather than the OLS-based test for heavy-tailed innovations.

4.2 Comparison with three unit root tests

We next conduct experiments to examine the performance of the proposed bootstrap method versus other commonly used unit root tests in the literature: (i) the bootstrap ADF coefficient test, (ii) the residual-based sieve bootstrap unit root test, S_n^* at Chang and Park (2003), and (iii) the wild bootstrap unit root test, MZ_α^b at Cavaliere and Taylor (2009a). We consider two trends functions, the constant trend and the linear trend, and the GLS method is employed to de-trend the data in the HB test as well as three other bootstrapping unit root tests.

The data generating process is

$$\Delta y_t = \phi y_{t-1} + u_t, \quad u_t = e_t + \pi e_{t-1},$$

where $\pi = -0.8$, -0.4, 0.0, 0.4, and 0.8 for different magnitudes of serial dependence, and $\phi = -c^*/n$ with $c^* = 0$ corresponding to the size and $c^* = 3.5$ or 7 to the local power. We consider two types of innovations for $\{e_t\}$: (i) *i.i.d.* standard normal random variables, and (ii) GARCH innovations as in (10). Furthermore, the perturbing sequence $\{\omega_t\}$ is generated from the Mammen's two-point distribution. Moreover, the MAIC at (6) is employed to select the order p in equation (2).

Table 3 presents the sizes of these four unit root tests for *i.i.d.* innovations of $\{e_t\}$. When $\pi = 0.8$, the HB test is sensitive while the other three tests are all conservative. For the case with $\pi = -0.8$, the wild bootstrap test is conservative, and the other three are sensitive. The HB test is able to control the sizes slightly better than the others although it has a more serious distortion at $\pi = -0.8$. In the case of GARCH innovations, Table 4 show similar findings to those in Table 3.

We further investigate the empirical powers of all four unit root tests under two types of innovations and two local alternatives. Tables 5-8 indicate that the HB test is almost uniformly superior to the ADF and sieve bootstrap tests. In addition, it is generally better than (or comparable to) the wild bootstrap test. In sum, HB performs well in the comparison with the other three tests.

5 Conclusion

In this paper, we propose the hybrid bootstrap method for unit root tests via the OLS and LAD estimators. We also obtain asymptotic distributions of the resulting tests, which are not only simple to use but also more powerful than traditional tests. Our proposed method could be applied to the unit root tests via the robust M estimators (see Lucas 1995; Ng and Perron 2001). In addition, it could be considered for testing cointegration (e.g., see Maddala and Kim 1998). Moreover, another useful extension of the LAD-based unit root test would involve allowing MA innovations as well as adding the constant trend or the linear trend into the model. We believe these efforts would further enhance the usefulness of the hybrid bootstrap unit root tests in data analysis.

Acknowledgment

We thank the editor, an associate editor and one anonymous referee for their valuable comments that led to a substantial improvement of the paper, and Hong Kong GRF grant HKU703710P for partial support.

Appendix: Proofs of Theorems 1-3

Proof of Theorem 1. For the sake of simplicity, we assume that $y_t = 0$ and $\Delta y_t = 0$ for $t \leq 0$, and the notations E^*, $O_p^*(1)$ and $o_p^*(1)$ correspond to the bootstrapped probability space. Let

$$A_{1n} = \sum_{t=1}^{n} y_{t-1}^* e_{t,p}^* - (\sum_{t=1}^{n} y_{t-1}^* x_{t,p}')(\sum_{t=1}^{n} x_{t,p} x_{t,p}')^{-1}(\sum_{t=1}^{n} x_{t,p} e_{t,p}^*),$$

$$A_{2n} = \sum_{t=1}^{n} \omega_t y_{t-1}^* e_{t,p}^* - (\sum_{t=1}^{n} \omega_t y_{t-1}^* x_{t,p}')(\sum_{t=1}^{n} \omega_t x_{t,p} x_{t,p}')^{-1}(\sum_{t=1}^{n} \omega_t x_{t,p} e_{t,p}^*),$$

$$B_{1n} = \sum_{t=1}^{n} y_{t-1}^{*2} - (\sum_{t=1}^{n} y_{t-1}^* x_{t,p}')(\sum_{t=1}^{n} x_{t,p} x_{t,p}')^{-1}(\sum_{t=1}^{n} y_{t-1}^* x_{t,p}),$$

and

$$B_{2n} = \sum_{t=1}^{n} \omega_t y_{t-1}^{*2} - (\sum_{t=1}^{n} \omega_t y_{t-1}^* x_{t,p}')(\sum_{t=1}^{n} \omega_t x_{t,p} x_{t,p}')^{-1}(\sum_{t=1}^{n} \omega_t y_{t-1}^* x_{t,p}),$$

where $y^*_{[n\tau]} = (1 - \sum_{i=1}^p \widehat{\psi}_i)^{-1} \sum_{i=1}^{[n\tau]} (\omega_i - 1)\widehat{e}_{i,p} + O_p(1)$, $e^*_{t,p} = e_{t,p} - (\widehat{\phi}_n - \phi)y_{t-1}$, and $x_{t,p} = (\Delta y_{t-1}, ..., \Delta y_{t-p})'$. As a result, $\widehat{\phi}^*_{1n} = A_{1n}/B_{1n}$, and $\widehat{\phi}^*_{2n} = A_{2n}/B_{2n}$.

We first show that, conditional on $y_1, ..., y_n$,

$$\|(n^{-1}\sum_{t=1}^n x_{t,p}x'_{t,p})^{-1}\| = O^*_p(1), \quad \|(n^{-1}\sum_{t=1}^n \omega_t x_{t,p}x'_{t,p})^{-1}\| = O^*_p(1), \tag{11}$$

$$\|\sum_{t=1}^n y^*_{t-1}x_{t,p}\| = O^*_p(np^{1/2}), \quad \|\sum_{t=1}^n \omega_t y^*_{t-1}x_{t,p}\| = O^*_p(np^{1/2}), \tag{12}$$

$$\|\sum_{t=1}^n x_{t,p}e^*_{t,p}\| = o^*_p(np^{-1/2}), \quad \|\sum_{t=1}^n \omega_t x_{t,p}e^*_{t,p}\| = o^*_p(np^{-1/2}), \tag{13}$$

and

$$\sum_{t=1}^n (\omega_t - 1)y^{*2}_{t-1} = o^*_p(n^2), \tag{14}$$

in probability. By Lemma 3.2 (a) in Chang and Park (2002), we have that

$$\|(\frac{1}{n}\sum_{t=1}^n x_{t,p}x'_{t,p})^{-1}\| = O_p(1). \tag{15}$$

In addition, for each $1 \le i, j \le p$, it is easy to see that

$$E^*[\sum_{t=1}^n (\omega_t - 1)\Delta y_{t-i}\Delta y_{t-j}]^2 = \sum_{t=1}^n \Delta y^2_{t-i}\Delta y^2_{t-j} = O_p(n).$$

This, together with Assumption 2, leads to

$$E^*\|\frac{1}{n}\sum_{t=1}^n \omega_t x_{t,p}x'_{t,p} - \frac{1}{n}\sum_{t=1}^n x_{t,p}x'_{t,p}\|^2 = O_p(n^{-1}p^2) = o_p(1). \tag{16}$$

In addition,

$$\left| \|(\frac{1}{n}\sum_{t=1}^n \omega_t x_{t,p}x'_{t,p})^{-1}\| - \|(\frac{1}{n}\sum_{t=1}^n x_{t,p}x'_{t,p})^{-1}\| \right| \le \|(\frac{1}{n}\sum_{t=1}^n \omega_t x_{t,p}x'_{t,p})^{-1} - (\frac{1}{n}\sum_{t=1}^n x_{t,p}x'_{t,p})^{-1}\|.$$

By (15)-(16) and using a method similar to Lemma 3 of Berk (1974), we are able to show that

$$\|(\frac{1}{n}\sum_{t=1}^n \omega_t x_{t,p}x'_{t,p})^{-1}\| - \|(\frac{1}{n}\sum_{t=1}^n x_{t,p}x'_{t,p})^{-1}\| = o^*_p(1).$$

This completes the proof of equation (11).

By Doob's inequality (see Hall and Heyde 1980) and Lemma 3.3 in Chang and Park (2002), we obtain that $\max_{1\le j\le n} |\sum_{t=1}^j \Delta y_t| = O_p(n^{1/2})$ and $n^{-1}\sum_{t=1}^n \widehat{e}^2_{t,p} = \sigma^2 + o_p(1)$, respectively. Then, for each $1 \le i \le p$, we are able to show that

$$E^*(\sum_{t=1}^{n} z_{t-1}^* \Delta y_{t-i})^2 = E^* \left[\sum_{j=1}^{n} (\omega_j - 1)\widehat{e}_{j,p}(\sum_{t=1}^{n} \Delta y_{t-i} - \sum_{t=1}^{j} \Delta y_{t-i})\right]^2$$
$$\le 4 \max_{1\le j\le n} |\sum_{t=1}^{j} \Delta y_t|^2 \cdot \sum_{j=1}^{n} \widehat{e}_{j,p}^2 = O_p(n^2)$$

and

$$E^*[\sum_{t=1}^{n}(\omega_t-1)z_{t-1}^*\Delta y_{t-i}]^2 = \sum_{t=1}^{n}(\Delta y_{t-i})^2 E^*(z_{t-1}^*)^2 \le \left[\sum_{t=1}^{n}(\Delta y_{t-i})^2\right]\cdot\left[\sum_{j=1}^{n}\widehat{e}_{j,p}^2\right] = O_p(n^2),$$

where $z_t^* = (1 - \sum_{i=1}^{p} \widehat{\psi}_i)y_t^* = \sum_{i=1}^{t}(\omega_i - 1)\widehat{e}_{i,p}$. The above results lead to equation (12).

Under H_0, $\max_{1\le j\le n} |y_j| = O_p(n^{1/2})$ (see Li and Li 2009). As a result, for $1 \le i \le p$, we have

$$|(\widehat{\phi}_n - \phi)\sum_{t=1}^{n} \Delta y_{t-i}y_{t-1}| \le |\widehat{\phi}_n - \phi| \cdot \max_{1\le j\le n} |y_j| \cdot \sum_{t=1}^{n} |\Delta y_{t-i}| = O_p(n^{1/2}).$$

In addition, under H_1,

$$|(\widehat{\phi}_n - \phi)\sum_{t=1}^{n} \Delta y_{t-i}y_{t-1}| \le |\widehat{\phi}_n - \phi| \cdot \sum_{t=1}^{n} |\Delta y_{t-i}y_{t-1}| = O_p(n^{1/2}).$$

The above results, together with Lemma 3.2 (c) of Chang and Park (2002) and Assumption 2, yields

$$\|\sum_{t=1}^{n} x_{t,p}e_{t,p}^*\| \le \|\sum_{t=1}^{n} x_{t,p}e_{t,p}\| + \|(\widehat{\phi}_n - \phi)\sum_{t=1}^{n} x_{t,p}y_{t-1}\| = o_p^*(np^{-1/2}). \tag{17}$$

Furthermore, for $1 \le i \le p$, it can be shown that

$$E^*[\sum_{t=1}^{n}(\omega_t - 1)\Delta y_{t-i}e_{t,p}]^2 = \sum_{t=1}^{n}(\Delta y_{t-i}e_{t,p})^2 = O_p(n) \tag{18}$$

and

$$E^*[(\widehat{\phi}_n - \phi)\sum_{t=1}^{n}(\omega_t - 1)\Delta y_{t-i}y_{t-1}]^2 = (\widehat{\phi}_n - \phi)^2\sum_{t=1}^{n}\Delta y_{t-i}^2 y_{t-1}^2 = O_p(1). \tag{19}$$

where $1 \le i \le p$. By (17)-(19), we complete the proof of equation (13).

By Burkholder's inequalities (Hall and Heyde, 1980), we obtain that

$$E^*[\sum_{t=1}^{n}(\omega_t - 1)z_{t-1}^{*2}]^2 = \sum_{t=1}^{n} E^*(z_{t-1}^*)^4 \le C_1 n E^*[\sum_{j=1}^{n}(\omega_j - 1)^2\widehat{e}_{j,p}^2]^2$$
$$\le C_1C_2n^3(\frac{1}{n}\sum_{j=1}^{n}\widehat{e}_{j,p}^2)^2 = o_p(n^4),$$

where $z_t^* = (1-\sum_{i=1}^p \widehat{\psi}_i)y_t^* = \sum_{i=1}^t(\omega_i-1)\widehat{e}_{i,p}$, C_1 is constant and $C_2 = E(\omega_t-1)^4 < \infty$. Hence, equation (14) holds, and then we finish the proofs of (11)-(14).

By (11)-(14) and Assumption 2, we are able to demonstrate that

$$\frac{1}{n^2}B_{1n} = \frac{1}{n^2}B_{2n} + o_p^*(1) = \frac{1}{n^2}\sum_{t=1}^n y_{t-1}^{*2} + o_p^*(1),$$

and then

$$n(\widehat{\phi}_{2n}^* - \widehat{\phi}_{1n}^*) = \frac{n^{-1}\sum_{t=1}^n y_{t-1}^*(\omega_t-1)e_{t,p}^*}{n^{-2}\sum_{t=1}^n y_{t-1}^{*2}} + o_p^*(1). \tag{20}$$

It is true that, under H_0, $\widehat{\phi}_n - \phi = O_p(n^{-1})$ and $\max_{1\le j\le n}|y_j| = O_p(n^{1/2})$ (see Li and Li, 2009). In addition, under H_1, $\{y_t\}$ is stationary and $\widehat{\phi}_n - \phi = O_p(n^{-1/2})$. Thus, by Lemmas 3.3 and 3.4 in Chang and Park (2002),

$$\frac{1}{n}\sum_{t=1}^n(\widehat{e}_{t,p}-e_{t,p})^2 \le (\widehat{\phi}_n-\phi)^2\frac{1}{n}\sum_{t=1}^n y_{t-1}^2 + \|\widehat{\Psi}-\Psi\|^2\cdot\frac{1}{n}\sum_{t=1}^n\|x_{t,p}\|^2 = o_p(1)$$

and then

$$\begin{aligned} E^*[\frac{1}{n}\sum_{t=1}^n z_{t-1}^*(\omega_t-1)(e_{t,p}^*-\widehat{e}_{t,p})]^2 &= \frac{1}{n^2}\sum_{t=1}^n\sum_{i=1}^{t-1}(\widehat{e}_{t,p})^2(e_{t,p}^*-\widehat{e}_{t,p})^2 \\ &\le \frac{2}{n}\sum_{t=1}^n(\widehat{e}_{t,p})^2\cdot\left[\frac{1}{n}\sum_{t=1}^n(\widehat{e}_{t,p}-e_{t,p})^2 + (\widehat{\phi}_n-\phi)^2\frac{1}{n}\sum_{t=1}^n y_{t-1}^2\right] = o_p(1). \end{aligned} \tag{21}$$

where $z_t^* = (1-\sum_{i=1}^p\widehat{\psi}_i)y_t^* = \sum_{i=1}^t(\omega_i-1)\widehat{e}_{i,p}$. It is noteworthy that the quantity $\{(\omega_t-1)\widehat{e}_{t,p}, t\in Z^+\}$ is a martingale difference with respect to $\{\mathcal{F}_t^*, t\in Z^+\}$, where $\mathcal{F}_t^* = \sigma(\omega_t,...,\omega_1,e_n,e_{n-1},...)$. It holds that, for any τ and ϵ, $\widehat{\sigma}_n^{-2}\cdot n^{-1}\sum_{t=1}^{[n\tau]}\widehat{e}_{t,p}^2 = \tau + o_p(1)$ and $n^{-1}\sum_{t=1}^{[n\tau]}\widehat{e}_{t,p}^2E^*\{(\omega_t-1)^2I[(\omega_t-1)\widehat{e}_{t,p}\ge n^{1/2}\epsilon]\} = o_p(1)$. Thus, applying Theorem 18.2 of Billingsley (1999), we have that,

$$\frac{1}{\widehat{\sigma}_n\sqrt{n}}\sum_{t=1}^{[n\tau]}(\omega_t-1)\widehat{e}_{t,p} \Rightarrow B^*(\tau)$$

in probability, where $\widehat{\sigma}_n^2 = n^{-1}\sum_{t=1}^n\widehat{e}_{t,p}^2 = \sigma^2 + o_p(1)$. This, together with (20) and (21), completes the proof of Theorem 1. □

Proof of Theorem 2. We first demonstrate that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}[e_t, \mathrm{sgn}(e_t), \mathrm{sgn}(e_t)x_t']' \Rightarrow [W_1(\tau), W_2(\tau), W_3'(\tau)]' = \mathbf{W}(\tau), \tag{22}$$

where $\mathbf{W}(\tau)$ is a $(p+2)$-dimensional Brownian motion process with covariance matrix $\tau\Omega$, and the matrix Ω is defined in Theorem 2. Let

$$\zeta_t = \lambda'[e_t, \mathrm{sgn}(e_t), \mathrm{sgn}(e_t)x_t']' \quad \mathit{and} \quad T_i = n^{-1/2}\sum_{t=1}^{i}\zeta_t,$$

where λ is a $(p+2)$-dimensional constant vector with $\lambda'\lambda \neq 0$. It is noteworthy that $\{\zeta_t, t \in Z\}$ is a martingale difference sequence with respect to the filtration $\{\mathcal{F}_t, t \in Z\}$ and $E(\zeta_t^2) = \lambda'\Omega\lambda$, where $\mathcal{F}_t = \sigma(e_t, e_{t-1}, ...)$. Accordingly, both sequences $\{\zeta_t\}$ and $\{E({\zeta_t}^2|\mathcal{F}_{t-1})\}$ are strictly stationary and ergodic, and $E{T_n}^2 = \lambda'\Omega\lambda$. Then, it can be verified that

$$\frac{1}{n}\sum_{t=1}^{n}\frac{E({\zeta_t}^2|\mathcal{F}_{t-1})}{E{T_n}^2} \to 1 \tag{23}$$

almost surely, and, for any $\epsilon > 0$,

$$\frac{1}{n}\sum_{t=1}^{n}E[{\zeta_t}^2 I(\zeta_t \geq \sqrt{n\mathrm{var}(\zeta_t)\epsilon})] \to 0, \tag{24}$$

as $n \to \infty$. The invariance principle for martingales (Hall and Heyde, 1980), together with (23) and (24), implies that

$$T_{[n\tau]} = \frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}\zeta_t \Rightarrow W(\tau),$$

where $W(\tau)$ is a Brownian motion with variance $\tau\lambda'\Omega\lambda$. By Cramér's device, we complete the proof of (22).

Following the Beveridge-Nelson representation (Chang and Park, 2002, Remark 2.2), Theorem 2.2 in Kurtz and Protter (1991), and (22), we further have that

$$\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n}\sigma_t^{-1}x_tx_t' = E[\sigma_t^{-1}x_tx_t'] + o_p(1), \qquad \frac{1}{\sqrt{n}}\sum_{t=1}^{n}x_t\,\mathrm{sgn}(e_t) \Rightarrow W_3(1),\\
&\frac{1}{n}\sum_{t=1}^{n}y_{t-1}\,\mathrm{sgn}(e_t) = \alpha\cdot\frac{1}{n}\sum_{t=1}^{n}z_{t-1}\,\mathrm{sgn}(e_t) + o_p(1) \Rightarrow \alpha\cdot\int W_1(\tau)dW_2(\tau),\\
&\frac{1}{n^2}\sum_{t=1}^{n}\sigma_t^{-1}y_{t-1}^2 = \alpha^2E(\sigma_t^{-1})\cdot\frac{1}{n^2}\sum_{t=1}^{n}z_{t-1}^2 + o_p(1) \Rightarrow \alpha^2E(\sigma_t^{-1})\cdot\int W_1^2(\tau)d\tau,
\end{aligned} \tag{25}$$

and

$$\frac{1}{n\sqrt{n}}\sum_{t=1}^{n}\sigma_t^{-1}y_{t-1}x_t = \alpha E(\sigma_t^{-1}x_t)\cdot\frac{1}{n\sqrt{n}}\sum_{t=1}^{n}z_{t-1} + o_p(1) \Rightarrow \alpha E(\sigma_t^{-1}x_t)\cdot\int W_1(\tau)d\tau, \tag{26}$$

where $\alpha = (1 - \sum_{i=1}^{p} \psi_i)^{-1}$ and $z_t = \sum_{i=1}^{t} e_i$.

Define the objective function

$$Q_n(\theta) = \sum_{t=p+2}^{n} |\Delta y_t - \phi y_{t-1} - \sum_{i=1}^{p} \psi_i \Delta y_{t-i}|.$$

Note that

$$|x-y| - |x| = -y\,\mathrm{sgn}(x) + 2\int_0^y I(x \le s) - I(x \le 0)ds,$$

for $x, y \in R$ and $x \neq 0$, where $\mathrm{sgn}(x)$ is equal to 1 for $x > 0$ and -1 for $x < 0$, see Knight (1998). Then, for any $v = (v_1, v_2')'$ with $v_1 \in R$ and $v_2 \in R^p$, we have that

$$\begin{aligned} &Q_n(v_1/n, \psi_0 + v_2/\sqrt{n}) - Q_n(0, \Psi_0) \\ &\qquad = \sum_{t=p+2}^{n} |e_t - \frac{v_1}{n} y_{t-1} - \frac{v_2'}{\sqrt{n}} x_t| - \sum_{t=p+2}^{n} |e_t| \\ &\qquad = -\frac{v_1}{n} \sum_{t=p+2}^{n} y_{t-1}\,\mathrm{sgn}(e_t) - \frac{v_2'}{\sqrt{n}} \sum_{t=p+2}^{n} x_t\,\mathrm{sgn}(e_t) + \xi_n, \end{aligned} \tag{27}$$

where $\theta_0 = (0, \Psi_0')'$ is the true parameter vector, $\iota_n(t) = n^{-1} v_1 y_{t-1} + n^{-1/2} v_2' x_t$, and

$$\xi_n = 2 \sum_{t=p+2}^{n} \int_0^{\iota_n(t)} I(e_t \le s) - I(e_t \le 0)ds.$$

Denote

$$\xi_{1n} = 2 \sum_{t=p+2}^{n} \int_0^{\iota_n(t)} F(s\sigma_t^{-1}) - F(0)ds \quad \text{and} \quad \xi_{2n} = 2 \sum_{t=p+2}^{n} \int_0^{\iota_n(t)} f(0)s\sigma_t^{-1}ds,$$

where $f(\cdot)$ and $F(\cdot)$ are, respectively, the density and the cumulative distribution of ε_t. We next show that $\xi_n = \xi_{1n} + o_p(1)$ and $\xi_{1n} = \xi_{2n} + o_p(1)$.

Note that the quantity $\xi_n - \xi_{1n}$ is the summation of a martingale difference sequence with respect to the filtration $\{\mathcal{F}_t, t \in Z\}$. Then, for any $\delta > 0$,

$$0.25E(\xi_n - \xi_{1n})^2 \le \sum_{t=p+2}^{n} E\left\{\int_0^{\iota_n(t)} [I(e_t \le s) - I(e_t \le 0)]ds\right\}^2 = a_n(\delta) + b_n(\delta), \tag{28}$$

where

$$a_n(\delta) = \sum_{t=p+2}^{n} E\left\{\int_0^{\iota_n(t)} [I(\varepsilon_t \le s\sigma_t^{-1}) - I(\varepsilon_t \le 0)]ds I(|\iota_n(t)|\sigma_t^{-1} \le \delta)\right\}^2$$

and

$$b_n(\delta) = \sum_{t=p+2}^{n} E\left\{\int_0^{\iota_n(t)} [I(\varepsilon_t \le s\sigma_t^{-1}) - I(\varepsilon_t \le 0)]ds I(|\iota_n(t)|\sigma_t^{-1} > \delta)\right\}^2.$$

By Assumption 3, we obtain that there exists a constant $\pi_1 > 0$ such that the density $f(\cdot)$ is continuous on the set $[-\pi_1, \pi_1]$. Furthermore,

$$\int_0^y I(x \leq s) - I(x \leq 0)ds = (y-x)I(0 < x < y) + (x-y)I(y < x < 0).$$

Moreover, for $\delta < \pi_1$, we have that

$$\begin{aligned} a_n(\delta) &= \sum_{t=p+2}^{n} E\{[\iota_n(t) - \varepsilon_t\sigma_t]^2[I(0 < \varepsilon_t < \iota_n(t)\sigma_t^{-1}) + I(\iota_n(t)\sigma_t^{-1} < \varepsilon_t < 0]\} \\ &\leq \delta \cdot C_1 n E[\iota_n(t)]^2, \end{aligned}$$

and $b_n(\delta) \leq nE\{[\iota_n(t)]^2 I(|\iota_n(t)|\sigma_t^{-1} > \delta)\}$, where $C_1 = \sup_{|x|\leq\pi_1} f(x)$ and

$$nE[\iota_n(t)]^2 \leq 2v_1^2 n^{-1}E(y_{t-1}^2) + 2v_2' E(x_t x_t')v_2 < \infty.$$

Thus, for a fixed δ, $b_n(\delta) \to 0$ as $n \to \infty$. Let $\delta \to 0$, we further obtain $a_n(\delta) \to 0$. These results, together with (28), imply that $0.25E(\xi_n - \xi_{1n})^2 = o(1)$. Consequently, $\xi_n = \xi_{1n} + o_p(1)$. Analogously, we can show that $\xi_{1n} = \xi_{2n} + o_p(1)$; and it is noteworthy that

$$\xi_{2n} = f(0)v' \begin{pmatrix} n^{-2}\sum_{t=p+2}^n \sigma_t^{-1} y_{t-1}^2 & n^{-3/2}\sum_{t=p+2}^n \sigma_t^{-1} y_{t-1}x_t' \\ n^{-3/2}\sum_{t=p+2}^n \sigma_t^{-1} y_{t-1}x_t & n^{-1}\sum_{t=p+2}^n \sigma_t^{-1} x_t x_t' \end{pmatrix} v.$$

The above results, in conjunction with (25), (26) and (27), lead to

$$\begin{aligned} &Q_n(v_1/n, \psi_0 + v_2/\sqrt{n}) - Q_n(0, \Psi_0) \\ &\Rightarrow -v' \begin{pmatrix} \alpha \int W_1(\tau)dW_2(\tau) \\ W_3(1) \end{pmatrix} + f(0)v' \begin{pmatrix} \alpha^2\Sigma_0 \int W_1^2(\tau)d\tau & \alpha\Sigma_1 \int W_1(\tau)d\tau \\ \alpha\Sigma_1 \int W_1(\tau)d\tau & \Sigma_2 \end{pmatrix} v, \end{aligned}$$

where $Q_n(v_1/n, \psi_0 + v_2/\sqrt{n})$ is a convex function with respect to v. Thus, by Knight (1998), we have that

$$\begin{pmatrix} n\widetilde{\phi}_n \\ \sqrt{n}(\widetilde{\Psi}_n - \psi_0) \end{pmatrix} \Rightarrow \frac{1}{2f(0)} \begin{pmatrix} \alpha^2\Sigma_0 \int W_1^2(\tau)d\tau & \alpha\Sigma_1 \int W_1(\tau)d\tau \\ \alpha\Sigma_1 \int W_1(\tau)d\tau & \Sigma_2 \end{pmatrix}^{-1} \begin{pmatrix} \alpha \int W_1(\tau)dW_2(\tau) \\ W_3(1) \end{pmatrix},$$

where $\widetilde{\Psi}_n = (\widetilde{\psi}_1, ..., \widetilde{\psi}_p)'$. After algebraic simplification with the fact that $(1 - \sum_{i=1}^p \widetilde{\psi}_i)^{-1} = \alpha + o_p(1)$, we complete the proof. □

Proof of Theorem 3. Consider the objective function

$$Q_n^*(\theta) = \sum_{t=p+2}^{n} |\Delta y_t - \phi y_{t-1}^* - \sum_{i=1}^{p} \psi_i \Delta y_{t-i} - \widetilde{\phi}_n y_{t-1}|.$$

Then, for any $v = (v_1, v_2')'$ with $v_1 \in R$ and $v_2 \in R^p$, we have that

$$\begin{aligned}
&Q_n^*(v_1/n, \psi_0 + v_2/\sqrt{n}) - Q_n^*(0, \Psi_0) \\
&\qquad = \left(\sum_{t=p+2}^{n} |e_t - \frac{v_1}{n} y_{t-1}^* - \frac{v_2'}{\sqrt{n}} x_t - \widetilde{\phi}_n y_{t-1}| - \sum_{t=p+2}^{n} |e_t| \right) \\
&\qquad\quad - \left(\sum_{t=p+2}^{n} |e_t - \widetilde{\phi}_n y_{t-1}| - \sum_{t=p+2}^{n} |e_t| \right). \qquad (29)
\end{aligned}$$

Applying a method similar to that in the proof of Theorem 2, we obtain that

$$\begin{aligned}
&\sum_{t=p+2}^{n} |e_t - \frac{v_1}{n} y_{t-1}^* - \frac{v_2'}{\sqrt{n}} x_t - \widetilde{\phi}_n y_{t-1}| - \sum_{t=p+2}^{n} |e_t| \\
&\quad = -\frac{v_1}{n} \sum_{t=p+2}^{n} y_{t-1}^* \operatorname{sgn}(e_t) - \frac{v_2'}{\sqrt{n}} \sum_{t=p+2}^{n} x_t \operatorname{sgn}(e_t) - \widetilde{\phi}_n \sum_{t=p+2}^{n} y_{t-1} \operatorname{sgn}(e_t) \\
&\qquad + 2f(0)\widetilde{\phi}_n \frac{v_1}{n} \sum_{t=p+2}^{n} \sigma_t^{-1} y_{t-1} y_{t-1}^* + 2f(0)\widetilde{\phi}_n \frac{v_2'}{\sqrt{n}} \sum_{t=p+2}^{n} \sigma_t^{-1} y_{t-1} x_t \\
&\qquad + f(0) v' \begin{pmatrix} n^{-2} \sum_{t=p+2}^{n} \sigma_t^{-1} y_{t-1}^{*2} & n^{-3/2} \sum_{t=p+2}^{n} \sigma_t^{-1} y_{t-1}^* x_t' \\ n^{-3/2} \sum_{t=p+2}^{n} \sigma_t^{-1} y_{t-1}^* x_t & n^{-1} \sum_{t=p+2}^{n} \sigma_t^{-1} x_t x_t' \end{pmatrix} v \\
&\qquad + f(0)\widetilde{\phi}_n^2 \sum_{t=p+2}^{n} \sigma_t^{-1} y_{t-1}^2 + o_p^*(1) \qquad (30)
\end{aligned}$$

and

$$\sum_{t=p+2}^{n} |e_t - \widetilde{\phi}_n y_{t-1}| - \sum_{t=p+2}^{n} |e_t| = -\widetilde{\phi}_n \sum_{t=p+2}^{n} y_{t-1} \operatorname{sgn}(e_t) + f(0)\widetilde{\phi}_n^2 \sum_{t=p+2}^{n} \sigma_t^{-1} y_{t-1}^2 + o_p^*(1). \quad (31)$$

Equations (29) to (31) imply that $\widetilde{\phi}_{1n}^* = 0.5 f^{-1}(0) \tilde{A}_{1n} / \tilde{B}_{1n} + o_p^*(n^{-1})$, where

$$\tilde{A}_{1n} = \sum_{t=1}^{n} y_{t-1}^* e_t^* - (\sum_{t=1}^{n} \sigma_t^{-1} y_{t-1}^* x_t')(\sum_{t=1}^{n} \sigma_t^{-1} x_t x_t')^{-1}(\sum_{t=1}^{n} x_t e_t^*),$$

$$\tilde{B}_{1n} = \sum_{t=1}^{n} \sigma_t^{-1} y_{t-1}^{*2} - (\sum_{t=1}^{n} \sigma_t^{-1} y_{t-1}^* x_t')(\sum_{t=1}^{n} \sigma_t^{-1} x_t x_t')^{-1}(\sum_{t=1}^{n} \sigma_t^{-1} y_{t-1}^* x_t),$$

and $e_t^* = \operatorname{sgn}(e_t) - 2f(0)\widetilde{\phi}_n \sigma_t^{-1} y_{t-1}$. Analogously, we can demonstrate that $\widetilde{\phi}_{2n}^* = 0.5 f^{-1}(0) \tilde{A}_{2n} / \tilde{B}_{2n} + o_p^*(n^{-1})$, where

$$\tilde{A}_{2n} = \sum_{t=1}^{n} \omega_t y_{t-1}^* e_t^* - (\sum_{t=1}^{n} \omega_t \sigma_t^{-1} y_{t-1}^* x_t')(\sum_{t=1}^{n} \omega_t \sigma_t^{-1} x_t x_t')^{-1}(\sum_{t=1}^{n} \omega_t x_t e_t^*),$$

and

$$\tilde{B}_{2n} = \sum_{t=1}^{n} \omega_t \sigma_t^{-1} y_{t-1}^{*2} - (\sum_{t=1}^{n} \omega_t \sigma_t^{-1} y_{t-1}^{*} x_t')(\sum_{t=1}^{n} \omega_t \sigma_t^{-1} x_t x_t')^{-1}(\sum_{t=1}^{n} \omega_t \sigma_t^{-1} y_{t-1}^{*} x_t).$$

Denote $\widetilde{\Sigma}_0 = n^{-1}\sum_{t=1}^{n} \sigma_t^{-1}$, $\widetilde{\Sigma}_1 = \sum_{t=1}^{n} \sigma_t^{-1} x_t$ and $\widetilde{\Sigma}_2 = \sum_{t=1}^{n} \sigma_t^{-1} x_t x_t'$. It can be shown that

$$\frac{1}{n}\sum_{t=1}^{n} \omega_t \sigma_t^{-1} x_t x_t' = \widetilde{\Sigma}_2 + o_p^*(1),$$
$$\frac{1}{n^2}\sum_{t=1}^{n} \omega_t \sigma_t^{-1} y_{t-1}^{*2} = \frac{1}{n^2}\sum_{t=1}^{n} \sigma_t^{-1} y_{t-1}^{*2} + o_p^*(1) = \widetilde{\Sigma}_0 \cdot \frac{1}{n^2}\sum_{t=1}^{n} y_{t-1}^{*2} + o_p^*(1)$$

and

$$\frac{1}{n\sqrt{n}}\sum_{t=1}^{n} \omega_t \sigma_t^{-1} y_{t-1}^{*} x_t = \frac{1}{n\sqrt{n}}\sum_{t=1}^{n} \sigma_t^{-1} y_{t-1}^{*} x_t + o_p^*(1) = \widetilde{\Sigma}_1 \cdot \frac{1}{n\sqrt{n}}\sum_{t=1}^{n} y_{t-1}^{*} + o_p^*(1).$$

As a result,

$$\frac{1}{n^2}\tilde{B}_{1n} = \frac{1}{n^2}\tilde{B}_{2n} + o_p^*(1) = \widetilde{\Sigma}_0 \cdot \frac{1}{n^2}\sum_{t=1}^{n} y_{t-1}^{*2} - \widetilde{\Sigma}_1'\widetilde{\Sigma}_2^{-1}\widetilde{\Sigma}_1 \cdot \left(\frac{1}{n\sqrt{n}}\sum_{t=1}^{n} y_{t-1}^{*}\right)^2 + o_p^*(1).$$

Subsequently, it can be demonstrated that

$$\frac{1}{n}\sum_{t=1}^{n} y_{t-1}^{*}(\omega_t - 1)e_t^{*} = \frac{1}{n}\sum_{t=1}^{n} y_{t-1}^{*}(\omega_t - 1)\operatorname{sgn}(e_t) + o_p^*(1),$$

and

$$\frac{1}{n}\sum_{t=1}^{n} (\omega_t - 1)e_t^{*} x_t = \frac{1}{n}\sum_{t=1}^{n} (\omega_t - 1)\operatorname{sgn}(e_t) x_t + o_p^*(1).$$

Thus,

$$n(\widetilde{\phi}_{2n}^{*} - \widetilde{\phi}_{1n}^{*}) = \frac{1}{2f(0)}\frac{\tilde{D}_n}{\widetilde{\Sigma}_0(n^{-2}\sum_{t=1}^{n} y_{t-1}^{*2}) - \widetilde{\Sigma}_1'\widetilde{\Sigma}_2^{-1}\widetilde{\Sigma}_1(n^{-3/2}\sum_{t=1}^{n} y_{t-1}^{*})^2} + o_p^*(1), \quad (32)$$

where

$$\tilde{D}_n = n^{-1}\sum_{t=1}^{n} y_{t-1}^{*}(\omega_t - 1)\operatorname{sgn}(e_t) - \widetilde{\Sigma}_1'\widetilde{\Sigma}_2^{-1}[n^{-3/2}\sum_{t=1}^{n} y_{t-1}^{*}][n^{-1}\sum_{t=1}^{n}(\omega_t - 1)\operatorname{sgn}(e_t)x_t].$$

Applying similar techniques to those used in the proof of Theorem 2, we can show that, conditional on $y_1, ..., y_n$,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]} (\omega_t - 1)[\widetilde{e}_t, \operatorname{sgn}(e_t), \operatorname{sgn}(e_t)x_t']' \Rightarrow [W_1^*(\tau), W_2^*(\tau), W_3^{*\prime}(\tau)]' = \mathbf{W}^*(\tau)$$

in probability, where $\mathbf{W}^*(\tau)$ is a $(p+2)$-dimensional Brownian motion process with covariance matrix $\tau\Omega$. Note that $L_n^* = n(\widetilde{\phi}_{2n}^{*} - \widetilde{\phi}_{1n}^{*})/(1 - \sum_{i=1}^{p}\widehat{\psi}_i)$. This, together with equation (32) and Theorem 2.2 of Kurtz and Protter (1991), completes the proof. □

References

Berk, K. N. (1974). Consistent autoregressive spectral estimates. *The Annals of Statistics 2*, 489–502.

Billingsley, P. (1999). *Convergence of Probability Measures* (2nd ed.). New York: Wiley.

Bollerslev, T. (1986). Generalized autoregression conditional heteroscedasticity. *Journal of Econometrics 31*, 307–327.

Cavaliere, G. and A. M. R. Taylor (2009a). Bootstrap M unit root tests. *Econometric Reviews 28*, 393–421.

Cavaliere, G. and A. M. R. Taylor (2009b). Heteroskedastic time series with a unit root. *Econometric Theory 25*, 1228–1276.

Chang, Y. and J. Y. Park (2002). On the asymptotic of ADF tests for unit roots. *Econometric Reviews 21*, 431–447.

Chang, Y. and J. Y. Park (2003). A sieve bootstrap for the test of a unit root. *Journal of Time Series Analysis 24*, 379–400.

Chatterjee, S. and A. Bose (2005). Generalized bootstrap for estimating equations. *The Annals of Statistics 33*, 414–436.

Chen, K., Z. Ying, H. Zhang, and L. Zhao (2008). Analysis of least absolute deviation. *Biometrika 95*, 107–122.

Dickey, D. A. (1976). *Estimation and hypothesis testing in nonstationary time series.* Ph. D. thesis, Iowa State University, Ames, Iowa.

Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association 74*, 427–431.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics 7*, 1–26.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1996). Efficient tests for an autoregressive unit root. *Econometrica 64*, 813–836.

Engle, R. F. (1982). Autoregression conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica 50*, 987–1008.

Engle, R. F. and G. Gonzalez-Rivera (1991). Semiparametric ARCH models. *Journal of Business and Economic Statistics 9*, 345–359.

Fuller, W. A. (1996). *Introduction to Statistical Time Series* (2 ed.). New York: Wiley.

Goncalves, S. and L. Kilian (2007). Asymptotic and bootstrap inference for AR(∞) processes with conditional heteroscedasticity. *Econometric Reviews 26*, 609–641.

Hall, P. and C. C. Heyde (1980). *Martingale Limit Theory and Its Application.* New York: Academic Press.

Herce, M. A. (1996). Asymptotic theory of LAD estimation in a unit root process with finite variance errors. *Econometric Theory 12*, 129–153.

Jin, Z., Z. Ying, and L. J. Wei (2001). A simple resampling method by perturbing the minimand. *Biometrika 88*, 381–390.

Knight, K. (1998). Limiting distributions for l_1 regression estimators under general conditions. *The Annals of Statistics 26*, 755–770.

Kurtz, T. G. and P. Protter (1991). Weak limit theorems to stochastic integrals and stochastic differential equations. *The Annals of Probability 19*, 1035–1070.

Li, G. and W. K. Li (2008). Least absolute deviation estimation for fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity. *Biometrika 95*, 399–414.

Li, G. and W. K. Li (2009). Least absolute deviation estimation for unit root processes with GARCH errors. *Econometric Theory 25*, 1208–1227.

Li, G. and W. K. Li (2011). Testing a linear time series models against its threshold extension. *Biometrika 98*, 243–250.

Ling, S. and W. K. Li (2003). Asymptotic inference for unit root processes with GARCH(1,1) errors. *Econometric Theory 19*, 541–564.

Liu, R. Y. (1988). Bootstrap procedures under some non-iid models. *The Annals of Statistics 16*, 1696–1708.

Lucas, A. (1995). Unit root tests based on M estimators. *Econometric Theory 11*, 331–346.

Maddala, G. S. and I.-M. Kim (1998). *Unit Roots, Cointegration, and Structural Change.* Cambridge: Cambridge University Press.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *The Annals of Statistics 21*, 255–285.

Moreno, M. and J. Romo (2000). Bootstrap tests for unit roots based on LAD estimation. *Journal of Statistical Planning and Inference 83*, 347–367.

Ng, S. and P. Perron (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica 69*, 1519–1554.

Palm, F. C., S. Smeekes, and J.-P. Urbain (2008). Bootstrap unit-root tests: comparison and extensions. *Journal of Time Series Analysis 29*, 371–401.

Paparoditis, E. and D. N. Politis (2003). Residual-based block bootstrap for unit root testing. *Econometrica 71*, 813–855.

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

Peng, L. and Q. Yao (2003). Least absolute deviations estimation for ARCH and GARCH models. *Biometrika 90*, 967–975.

Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica 55*, 277–301.

Phillips, P. C. B. (2010). Bootstrapping I(1) data. *Journal of Econometrics 158*, 280–284.

Phillips, P. C. B. and P. Perron (1988). Testing for a unit root in time series regression. *Biometrika 75*, 335–346.

Said, E. S. and D. A. Dickey (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika 71*, 599–607.

Seo, B. (1999). Distribution theory for unit root tests with conditional heteroskedasticity. *Journal of Econometrics 91*, 113–144.

Table 1: Rejection rates of the test S_n^* under three sample sizes, three significance levels, four ϕ values, and four perturbing distributions.

	$n = 100$			$n = 200$			$n = 300$		
ϕ	1%	5%	10%	1%	5%	10%	1%	5%	10%
				Exponential distribution					
0.0	0.015	0.058	0.110	0.011	0.057	0.111	0.012	0.052	0.107
-0.01	0.025	0.087	0.162	0.038	0.139	0.262	0.050	0.188	0.317
-0.025	0.045	0.159	0.271	0.097	0.337	0.554	0.183	0.539	0.756
-0.05	0.091	0.326	0.525	0.345	0.761	0.912	0.641	0.950	0.991
				Rademacher distribution					
0.0	0.009	0.047	0.099	0.007	0.050	0.096	0.009	0.050	0.096
-0.01	0.014	0.077	0.147	0.029	0.124	0.238	0.037	0.163	0.302
-0.025	0.030	0.125	0.231	0.074	0.310	0.518	0.155	0.499	0.747
-0.05	0.062	0.277	0.468	0.284	0.722	0.898	0.612	0.935	0.989
				Mammen's distribution					
0.0	0.011	0.054	0.101	0.009	0.054	0.102	0.011	0.052	0.102
-0.01	0.021	0.079	0.154	0.033	0.130	0.244	0.046	0.174	0.311
-0.025	0.037	0.139	0.255	0.088	0.311	0.522	0.174	0.512	0.747
-0.05	0.079	0.300	0.501	0.321	0.731	0.904	0.626	0.939	0.989
				Mixture distribution					
0.0	0.011	0.052	0.100	0.011	0.057	0.103	0.010	0.048	0.095
-0.01	0.017	0.079	0.154	0.033	0.128	0.245	0.041	0.170	0.313
-0.025	0.036	0.140	0.255	0.086	0.317	0.532	0.168	0.511	0.754
-0.05	0.074	0.295	0.500	0.313	0.743	0.903	0.629	0.941	0.990

Table 2: Rejection rates of the tests S_n^* and L_n^* for GARCH innovations with $t(3)$ innovations.

	$n = 100$			$n = 200$			$n = 300$		
ϕ	1%	5%	10%	1%	5%	10%	1%	5%	10%
				The S_n^* Test					
0.0	0.014	0.067	0.109	0.017	0.053	0.105	0.011	0.048	0.107
-0.01	0.019	0.095	0.165	0.046	0.149	0.269	0.036	0.184	0.320
-0.025	0.034	0.151	0.276	0.109	0.345	0.527	0.186	0.528	0.729
-0.05	0.091	0.327	0.526	0.328	0.705	0.879	0.593	0.890	0.970
				The L_n^* Test					
0.0	0.007	0.040	0.098	0.010	0.049	0.094	0.011	0.044	0.106
-0.01	0.012	0.075	0.160	0.035	0.164	0.321	0.042	0.254	0.491
-0.025	0.020	0.160	0.357	0.102	0.466	0.715	0.283	0.774	0.918
-0.05	0.071	0.407	0.653	0.435	0.874	0.963	0.836	0.988	0.996

Table 3: Sizes of the HB test S_n^* and three other tests for *i.i.d.* innovations $\{e_t\}$. The significance level is 5%, and $\pi = \pm 0.8$, ± 0.4 or 0.0.

	Constant trend					Linear trend				
n	−0.8	−0.4	0.0	0.4	0.8	−0.8	−0.4	0.0	0.4	0.8
					HB tests					
100	0.230	0.074	0.050	0.049	0.051	0.412	0.115	0.044	0.056	0.076
200	0.152	0.062	0.051	0.047	0.052	0.269	0.064	0.043	0.061	0.085
300	0.132	0.055	0.050	0.051	0.058	0.201	0.056	0.051	0.053	0.072
					ADF tests					
100	0.103	0.042	0.042	0.022	0.013	0.146	0.048	0.028	0.015	0.006
200	0.079	0.042	0.037	0.044	0.024	0.075	0.035	0.024	0.020	0.023
300	0.087	0.044	0.046	0.039	0.036	0.072	0.045	0.035	0.029	0.025
					Sieve bootstrap tests					
100	0.095	0.034	0.035	0.016	0.013	0.153	0.050	0.024	0.016	0.007
200	0.073	0.040	0.031	0.025	0.025	0.089	0.035	0.025	0.020	0.019
300	0.086	0.049	0.053	0.039	0.028	0.079	0.049	0.033	0.028	0.025
					Wild bootstrap tests					
100	0.036	0.030	0.039	0.029	0.017	0.035	0.030	0.040	0.028	0.013
200	0.024	0.033	0.046	0.048	0.024	0.024	0.027	0.036	0.035	0.027
300	0.015	0.045	0.052	0.042	0.036	0.019	0.042	0.042	0.044	0.035

Table 4: Sizes of the HB test S_n^* and three other tests for GARCH innovations $\{e_t\}$. The significance level is 5%, and $\pi = \pm 0.8$, ± 0.4 or 0.0.

	Constant trend					Linear trend				
n	−0.8	−0.4	0.0	0.4	0.8	−0.8	−0.4	0.0	0.4	0.8
					HB tests					
100	0.228	0.071	0.038	0.041	0.054	0.389	0.123	0.054	0.058	0.079
200	0.155	0.060	0.045	0.048	0.061	0.259	0.065	0.057	0.054	0.065
300	0.141	0.056	0.050	0.052	0.060	0.197	0.053	0.045	0.047	0.059
					ADF tests					
100	0.081	0.033	0.019	0.011	0.010	0.129	0.032	0.018	0.008	0.010
200	0.072	0.036	0.036	0.028	0.017	0.065	0.029	0.026	0.018	0.007
300	0.073	0.032	0.038	0.031	0.035	0.079	0.031	0.028	0.022	0.021
					Sieve bootstrap tests					
100	0.090	0.036	0.025	0.013	0.015	0.146	0.038	0.020	0.011	0.013
200	0.076	0.040	0.039	0.032	0.024	0.085	0.036	0.030	0.021	0.009
300	0.081	0.046	0.039	0.032	0.041	0.092	0.041	0.030	0.028	0.025
					Wild bootstrap tests					
100	0.038	0.028	0.039	0.025	0.021	0.033	0.030	0.041	0.021	0.018
200	0.022	0.037	0.044	0.042	0.031	0.026	0.041	0.043	0.031	0.017
300	0.020	0.034	0.047	0.044	0.040	0.016	0.042	0.045	0.039	0.028

Table 5: Powers of the HB test S_n^* and three other tests for *i.i.d.* innovations $\{e_t\}$ with $\phi = -3.5/n$. The significance level is 5%, and $\pi = \pm 0.8$, ± 0.4 or 0.0.

	Constant trend					Linear trend				
n	−0.8	−0.4	0.0	0.4	0.8	−0.8	−0.4	0.0	0.4	0.8
HB tests										
100	0.418	0.193	0.115	0.131	0.135	0.488	0.178	0.070	0.079	0.102
200	0.362	0.181	0.121	0.141	0.172	0.348	0.120	0.063	0.088	0.101
300	0.340	0.164	0.129	0.137	0.143	0.295	0.113	0.080	0.087	0.108
ADF tests										
100	0.167	0.105	0.085	0.067	0.043	0.178	0.063	0.031	0.019	0.013
200	0.185	0.123	0.096	0.097	0.087	0.097	0.061	0.044	0.049	0.010
300	0.200	0.124	0.116	0.103	0.070	0.105	0.066	0.059	0.042	0.030
Sieve bootstrap tests										
100	0.190	0.096	0.069	0.056	0.033	0.189	0.067	0.029	0.020	0.017
200	0.184	0.120	0.103	0.101	0.080	0.109	0.059	0.044	0.042	0.017
300	0.199	0.111	0.100	0.093	0.088	0.131	0.067	0.059	0.042	0.030
Wild bootstrap tests										
100	0.064	0.084	0.114	0.084	0.043	0.059	0.058	0.059	0.041	0.023
200	0.023	0.110	0.119	0.110	0.089	0.027	0.058	0.068	0.070	0.031
300	0.023	0.124	0.134	0.114	0.073	0.024	0.070	0.076	0.071	0.043

Table 6: Powers of the HB test S_n^* and three other tests for $i.i.d.$ innovations $\{e_t\}$ with $\phi = -7.0/n$. The significance level is 5%, and $\pi = \pm 0.8$, ± 0.4 or 0.0.

	Constant trend					Linear trend				
n	−0.8	−0.4	0.0	0.4	0.8	−0.8	−0.4	0.0	0.4	0.8
					HB tests					
100	0.600	0.359	0.231	0.214	0.247	0.639	0.243	0.138	0.103	0.163
200	0.540	0.334	0.276	0.248	0.275	0.493	0.213	0.128	0.156	0.187
300	0.535	0.323	0.279	0.277	0.266	0.447	0.200	0.140	0.148	0.194
					ADF tests					
100	0.274	0.203	0.178	0.130	0.080	0.248	0.088	0.082	0.031	0.019
200	0.283	0.239	0.227	0.167	0.137	0.149	0.121	0.093	0.068	0.048
300	0.330	0.243	0.249	0.221	0.151	0.174	0.130	0.101	0.080	0.062
					Sieve bootstrap tests					
100	0.301	0.191	0.184	0.122	0.060	0.263	0.085	0.077	0.030	0.022
200	0.309	0.251	0.197	0.175	0.144	0.177	0.121	0.088	0.072	0.044
300	0.370	0.274	0.207	0.198	0.180	0.200	0.126	0.099	0.076	0.069
					Wild bootstrap tests					
100	0.117	0.166	0.213	0.152	0.078	0.103	0.061	0.097	0.045	0.031
200	0.058	0.214	0.284	0.223	0.140	0.038	0.086	0.120	0.094	0.052
300	0.059	0.252	0.301	0.261	0.176	0.028	0.098	0.131	0.102	0.074

Table 7: Powers of the HB test S_n^* and three other tests for GARCH innovations $\{e_t\}$ with $\phi = -3.5/n$. The significance level is 5%, and $\pi = \pm 0.8$, ± 0.4 or 0.0.

	Constant trend					Linear trend				
n	−0.8	−0.4	0.0	0.4	0.8	−0.8	−0.4	0.0	0.4	0.8
					HB tests					
100	0.452	0.173	0.096	0.132	0.145	0.475	0.149	0.092	0.084	0.123
200	0.347	0.162	0.123	0.137	0.146	0.337	0.111	0.065	0.084	0.125
300	0.288	0.155	0.118	0.131	0.143	0.285	0.110	0.077	0.100	0.094
					ADF tests					
100	0.206	0.084	0.059	0.068	0.051	0.176	0.049	0.039	0.018	0.010
200	0.174	0.111	0.091	0.084	0.062	0.114	0.040	0.039	0.034	0.028
300	0.169	0.110	0.089	0.096	0.093	0.097	0.050	0.046	0.051	0.023
					Sieve bootstrap tests					
100	0.217	0.089	0.069	0.072	0.054	0.201	0.056	0.046	0.020	0.015
200	0.191	0.110	0.103	0.089	0.061	0.136	0.050	0.050	0.039	0.033
300	0.182	0.122	0.099	0.099	0.098	0.119	0.068	0.059	0.063	0.027
					Wild bootstrap tests					
100	0.090	0.061	0.072	0.077	0.055	0.103	0.039	0.050	0.031	0.020
200	0.027	0.096	0.100	0.094	0.063	0.029	0.047	0.051	0.043	0.029
300	0.028	0.099	0.107	0.108	0.096	0.020	0.053	0.053	0.062	0.039

Table 8: Powers of the HB test S_n^* and three other tests for GARCH innovations $\{e_t\}$ with $\phi = -7.0/n$. The significance level is 5%, and $\pi = \pm 0.8$, ± 0.4 or 0.0.

	Constant trend					Linear trend				
n	-0.8	-0.4	0.0	0.4	0.8	-0.8	-0.4	0.0	0.4	0.8
					HB tests					
100	0.609	0.320	0.228	0.200	0.223	0.601	0.251	0.124	0.093	0.145
200	0.535	0.306	0.266	0.246	0.258	0.472	0.179	0.112	0.144	0.145
300	0.491	0.318	0.261	0.249	0.273	0.408	0.183	0.119	0.145	0.167
					ADF tests					
100	0.249	0.181	0.154	0.088	0.064	0.261	0.092	0.066	0.024	0.021
200	0.283	0.199	0.201	0.150	0.109	0.172	0.076	0.062	0.057	0.032
300	0.314	0.243	0.227	0.190	0.159	0.169	0.095	0.075	0.072	0.054
					Sieve bootstrap tests					
100	0.271	0.196	0.169	0.101	0.068	0.286	0.111	0.081	0.029	0.024
200	0.313	0.211	0.217	0.168	0.127	0.202	0.091	0.069	0.075	0.040
300	0.324	0.252	0.243	0.208	0.168	0.202	0.109	0.093	0.083	0.066
					Wild bootstrap tests					
100	0.106	0.144	0.180	0.111	0.056	0.129	0.063	0.096	0.049	0.027
200	0.058	0.170	0.225	0.165	0.133	0.043	0.062	0.091	0.085	0.037
300	0.045	0.232	0.252	0.223	0.183	0.025	0.070	0.105	0.101	0.068