Journal of Mathematical Research & Exposition
Jan., 2011, Vol. 31, No. 1, pp. 79–90
DOI:10.3770/j.issn:1000-341X.2011.01.009
Http://jmre.dlut.edu.cn

Modified Testing for Structural Changes in Autoregressive Processes

Hao JIN[1,2,*], **Zheng TIAN**[2,3,*], **Yun Feng YANG**[1]

1. *School of Science, Xi'an University of Science and Technology, Shaanxi* 710054*, P. R. China;*
2. *Department of Applied Mathematics, Northwestern Polytechnical University, Shaanxi* 710072*, P. R. China;*
3. *State Key Laboratory of Remote Sensing Science, Chinese Academy of Sciences, Beijing* 100101*, P. R. China*

Abstract In this paper, we consider the problem of detecting for structural changes in the autoregressive processes including AR(p) process. In performing a test, we employ the conventional residual CUSUM of squares test (RCUSQ) statistic. The RCUSQ test is based on the subsampling method introduced by Jach and Kokoszka [J. Methodology and Computing in Applied Probability 25(2004)]. It is shown that under regularity conditions, the asymptotic distribution of the test statistic is the function of a standard Brownian bridge. Simulation results as to AR(1) process and an example of real data analysis are provided for illustration.

Keywords subsampling; RCUSQ test; Brownian bridge; structural changes.

Document code A
MR(2010) Subject Classification 62F12; 60F17
Chinese Library Classification O212.9

1. Introduction

The problem of testing for structural changes in statistical models has been an important issue among both theoreticians and practitioners. Research into this problem originally began with iid sample; for a review of earlier work, see [3, 4, 9–11, 19, 20, 23]. Subsequently, the issue became very popular in the time series context since series often suffer from structural changes. Particularly, econometric time series exhibit changes in their underlying model because a myriad of political and economic factors can cause the relationships among economic variables to change over time. For references, see [1, 7, 13, 14, 16, 17, 24] and the papers cited therein.

Recently, Lee et al. [22] extended the CUSUM test to a more general case, motivated by the conjecture: given a parameter of interest and its consistent estimator, under what conditions can the estimator be utilized to detect a change in that parameter. The result of Lee et al. [22] indicates that the CUSUM test performs well for a broad class of stationary processes including linear processes.

Received March 23, 2009; Accepted May 28, 2010
Supported by the National Natural Science Foundation of China (Grant Nos. 70873094; 08&ZD046; 10926197; 2010039).
* Corresponding author
E-mail address: jinhao_2004@126.com (H. JIN); zhtian@nwpu.edu.cn (Z. TIAN)

In this article, we focus on the residual CUSUM of squares test (RCUSQ) based on the subsampling methodology [2, 6] in stochastic processes, which can conventionally discard correlation effects and enhance the performance of the test. Particularly, special attention is paid to the linear autoregressive (LAR) time series since they accommodate important linear time series models, such as the autoregressive processes AR(p), which have been central to the analysis of data with linear characteristics [5, 8, 18]. Hence, the objective of this paper is to derive the asymptotic distribution of RCUSQ test and extend the theory of Jach and Kokoszka [2] to test for structural changes in AR(p) process at some unknown date.

The organization of this paper is as follows. In Section 2, we present the regular conditions under which the RCUSQ test statistic converges weakly to the function of a standard Brownian bridge. In Section 3, as an illustration we consider the structural changes problem in AR(p) process. Simulation results related to AR(1) process and an empirical application are reported in Section 4. We provide brief concluding remarks in Section 5.

2. Assumptions and models

We consider the following models:

$$\begin{aligned} y_t &= \mu + \xi_t, \\ \xi_t &= \alpha_1 \xi_{t-1} + \alpha_2 \xi_{t-2} + \cdots + \alpha_p \xi_{t-p} + \varepsilon_t, \end{aligned} \tag{1}$$

where p is a finite positive integer and $\{\xi_t\}$ is an AR(p) process. Assume the innovations processes $\{\varepsilon_t\}$ satisfy $E\varepsilon_t = 0$ and $E\varepsilon_t^2 = \sigma^2$.

Denoting by $\theta = (\mu, \alpha_1, \ldots, \alpha_p, \sigma)$ the parameter vector in (1), we test the null hypothesis,

$$H_0 : y_1, \ldots, y_T \quad \text{is a sample for some} \quad \theta,$$

against the change-point alternative

$$H_1 : \exists\ \theta,\ \theta^*, \quad \text{satisfying} \quad \theta \neq \theta^*,$$

and such that the sample $y_1, \ldots, y_T$ has the form

$$y_t = \begin{cases} \mu + \xi_t, & \xi_t = \alpha_1 \xi_{t-1} + \alpha_2 \xi_{t-2} + \cdots + \alpha_p \xi_{t-p} + \varepsilon_t, \quad t \leq k^*; \\ \mu^* + \xi_t, & \xi_t = \alpha_1^* \xi_{t-1} + \alpha_2^* \xi_{t-2} + \cdots + \alpha_p^* \xi_{t-p} + \varepsilon_t^*, \quad t > k^*, \end{cases}$$

where $k^* = [T\tau^*]$, $0 < \tau^* < 1$ is unknown and fixed.

We state the assumptions which are needed to prove asymptotic validity of our approach.

Assumption 2.1 *Assume the innovations processes $\{\varepsilon_t\}$ are independent identical distribution and satisfy $E|\varepsilon_t|^{4+\delta} < \infty$, where $\delta > 0$.*

Assumption 2.2 *$\{\eta_t, \varepsilon_t\}$ is strong mixing.*

Assumption 2.3 *All the roots of $1 - \alpha_1 z - \alpha_2 z^2 - \cdots - \alpha_p z^p = 0$ lie out of the unit circle.*

Remark 2.1 Assumptions 2.1 and 2.2 are basic conditions to derive the asymptotic distribution of RCUSQ test. The last Assumption can guarantee process $\{\xi_t\}$ be stationary, viz., $\xi_t = \sum_{j=0}^{\infty} \varphi_j \varepsilon_{t-j}$. Beveridge and Nelson [21] decomposed $T^{-1}\sum_{t=1}^{T} \xi_t = T^{-1}\varphi(1)\sum_{t=1}^{T} \varepsilon_t + o_p(1)$, where $\varphi(1) = \sum_{j=0}^{\infty} \varphi_j < \infty$, which shows that the rate of convergence for $T^{-1}\sum_{t=1}^{T} \xi_t$ and $T^{-1}\sum_{t=1}^{T} \varepsilon_t$ are the same.

Our approach also relies on the following results.

Lemma 2.1 ([12]) *If Assumption 2.1 holds, then*

$$T^{1/2}(\hat{\alpha}_i - \alpha_i) \text{ has a proper, nondegenerate limiting distribution,}$$

where $\hat{\alpha}_i$ is ordinary least estimators.

Remark 2.2 The results indicate that $\hat{\alpha}_i - \alpha_i = O_p(T^{-1/2})$. Since series $\{\varepsilon_t\}$ are not observable, our test is based on residuals $\hat{\varepsilon}_t^2$ instead of ε_t^2, which are obtained via estimating the unknown parameters, and these estimators play an important role to detect changes.

3. Asymptotic under the null hypothesis

In this section, we derive the asymptotic distribution of RCUSQ test under null hypothesis.

Theorem 3.1 *Assume the Assumptions 2.1–2.3 hold, let*

$$\Xi_T = \max_{q+1\le k\le T} \sqrt{T}\Big|\frac{\sum_{t=q+1}^{k} \hat{\varepsilon}_t^2}{\sum_{t=q+1}^{T} \hat{\varepsilon}_t^2} - \frac{k-q}{T-q}\Big|.$$

Then under null hypothesis,

$$\Xi_T \xrightarrow{P} \frac{\tau}{\sigma^2} \max_{0\le v\le 1} |BB(v)|,$$

where $\tau^2 = E\varepsilon_t^4 - \sigma^4$ and $v = k/T$. $BB(v)$ is a standard Brownian bridge.

Remark 3.1 In practice, we replace the unknown parameters τ^2 and σ^2 by $\hat{\tau}^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_t^4 - \hat{\sigma}^4$ and $\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_t^2$, respectively, and set $q = [(\log T)^2]$. When there are no changes, the residuals $\hat{\varepsilon}_t$ can behave like ε_t and construct the RCUSQ test. However, the test statistic is conservative to reject null hypothesis and produces low power even for large samples. In order to overcome the problem, we now resort to the subsampling methodology. Our goal is to develop an approximation to the null distribution of Ξ_T.

To describe the idea, the steps are constructed as follows:

Step 1. Compute the residuals from the regression of y_t on an intercept: $\hat{\xi}_t = y_t - \hat{\mu}$, $\hat{\mu} = T^{-1}\sum_{j=1}^{T} y_j$, $t = 1, 2, \ldots, T$.

Step 2. Compute the ordinary least estimator of parameters $\hat{\alpha}_1, \ldots, \hat{\alpha}_p$ based on $\hat{\xi}_1, \ldots, \hat{\xi}_T$.

Step 3. Compute the estimator of innovations $\hat{\varepsilon}_t = \hat{\xi}_t - \hat{\alpha}_1\hat{\xi}_{t-1} - \hat{\alpha}_2\hat{\xi}_{t-2} - \cdots - \hat{\alpha}_p\hat{\xi}_{t-p}$, $t = 1, 2, \ldots, T$.

Step 4. Compute the centered residuals $\varepsilon_t^0 = \hat{\varepsilon}_t - T^{-1}\sum_{j=1}^{T} \hat{\varepsilon}_j$, $1 \le t \le T$.

Step 5. Set $\varepsilon_i^1 = \varepsilon_{i+q}^0$, $i = 1, \ldots, T-q$. Fixed an integer $b < T$ and construct $T-b-q$ processes of length b which satisfy null hypothesis. For $l = 1, \ldots, T-b-q$, the l-th process is defined by: $y_0(l) = 0, y_1(l) = \varepsilon_l^1, y_2(l) = \varepsilon_{l+1}^1, \ldots, y_b(l) = \varepsilon_{l+b-1}^1$.

Step 6. Analogous to Ξ_T, we construct $\Xi^1_{b,l}$ based on $y_0(l), y_1(l), \ldots, y_b(l)$. Denote by $\Xi^1_b(\alpha)$ the $(1-\alpha)$-th quantile of the empirical distribution of the $T-b-q$ value $\Xi^1_{b,l}$. We reject null hypothesis if $\Xi_T > \Xi^1_b(\alpha)$, because the empirical distribution of $\Xi^1_{b,l}$ is an approximation to the sampling distribution of Ξ_T under null hypothesis.

We construct $\Xi^1_{b,l}$ by:

$$\Xi^1_{b,l} = \max_{0\le v\le 1} \sqrt{b}\Big|\frac{\sum_{i=l}^{l+[(b-1)v]} \varepsilon_i^{12} - v\sum_{i=l}^{l+(b-1)} \varepsilon_i^{12}}{\sum_{i=l}^{l+(b-1)} \varepsilon_i^{12}}\Big|.$$

Denote

$$\tilde{G}_b(x) = \frac{1}{T-b-q}\sum_{l=1}^{T-b-q} I\{\Xi^1_{b,l} \le x\}$$

and

$$G(x) = P\Big(\frac{\tau}{\sigma^2}\max_{0\le v\le 1}|BB(v)| \le x\Big).$$

Theorem 3.2 *Assume all the Assumptions hold, $b\to\infty$ and $b/T\to 0$ as $T\to\infty$. Then for every $x>0$,*

$$\tilde{G}_b(x) \xrightarrow{P} G(x).$$

Remark 3.3 Theorem 3.2 implies that the subsampling test has asymptotically correct size, and its proof is in Appendix A.

4. Simulation and empirical application

4.1 Simulation

In the section, we present the results of a simulation study intended to assess the performance of the subsampling test procedures in Section 3. The independent identically distributioned random variables $\{\varepsilon_t\}$ are $N(0,\sigma^2)$. All simulations were based on 5000 replications. We report empirical rejection frequencies of the tests with $T=(500,800,1000)$ for tests run at 5% critical value in various combinations. Since the tests are affected by a choice of b for fairly large samples, subsample size b is considered approximately equal to anything between 15% and 20% of the sample size T.

We consider the following models:

$$y_t = \mu + \xi_t, \quad \xi_t = \alpha_1\xi_{t-1} + \varepsilon_t.$$

The empirical sizes are calculated from the AR(1) model with $\sigma=1$ and $(\mu,\alpha_1)=(2,0.1)$, $(3,0.1)$, $(5,0.1)$. The figures in Table 1 indicate the proportion of the number of rejections of null hypothesis H_0 under which no structural changes are assumed to occur.

However, when studying the power property of test, we consider the same model above, allowing a change in μ, α_1 and σ simultaneously. The Data Generating Processes satisfy

$$y_t = \begin{cases} \mu + \xi_t, & \xi_t = \alpha_1\xi_{t-1} + \varepsilon_t, \quad t\le k^*; \\ \mu^* + \xi_t, & \xi_t = \alpha_1^*\xi_{t-1} + \varepsilon_t^*, \quad t > k^*, \end{cases}$$

where ε_t^* are IIDN$(0,(\sigma^*)^2)$, $k^*=[0.5T]$ (in bracket) and $k^*=[0.75T]$ (out bracket), denoted by $\kappa=(\sigma^*/\sigma)^2$.

As proposed by Jach and Kokoszka [2], a finite sample correction is introduced. They reject null hypothesis if $\Xi_T > \Xi_b^1(\alpha-c)$, where $c=c(T,\alpha,b)$ is a function of T (sample size), α (significance level) and b (subsample size). The following formula for c gives satisfactory result:

$$c=\frac{\sqrt{b\alpha}}{2\sqrt{1.6T}}.$$

For $T=500,800,1000$, we consider the two cases: (i) $\mu^*=3,5$, $\kappa=1.2,1.3,1.5$, $\alpha_1=0.1$ and $\alpha_1^*=\pm0.3$; (ii) $\mu^*=3,5$, $\kappa=1.2,1.3,1.5$, $\alpha_1=0.1$ and $\alpha_1^*=\pm0.5$.

α_1	T	b	$\mu=2$	$\mu=3$	$\mu=5$
0.1	500	80	0.044(0.046)	0.051(0.049)	0.047(0.046)
	800	130	0.050(0.047)	0.042(0.045)	0.048(0.049)
	1000	185	0.051(0.048)	0.052(0.046)	0.053(0.048)

Table 1 Empirical sizes of subsampling test

				$\mu^*=3$			$\mu^*=5$		
α_1	α_1^*	T	b	$\kappa=1.2$	$\kappa=1.3$	$\kappa=1.5$	$\kappa=1.2$	$\kappa=1.3$	$\kappa=1.5$
0.1	0.3	500	80	0.241(0.246)	0.448(0.453)	0.833(0.826)	0.273(0.266)	0.472(0.467)	0.867(0.872)
		800	130	0.377(0.373)	0.667(0.653)	0.962(0.958)	0.395(0.401)	0.694(0.692)	0.971(0.975)
		1000	185	0.437(0.434)	0.751(0.749)	0.989(0.994)	0.458(0.466)	0.776(0.773)	1.000(1.000)
0.1	-0.3	500	80	0.237(0.238)	0.443(0.452)	0.827(0.833)	0.285(0.278)	0.477(0.481)	0.874(0.885)
		800	130	0.384(0.378)	0.658(0.661)	0.955(0.968)	0.397(0.398)	0.690(0.687)	0.966(0.968)
		1000	185	0.432(0.428)	0.754(0.755)	0.991(0.997)	0.462(0.453)	0.784(0.779)	0.999(1.000)

Table 2 Empirical powers of subsampling test

				$\mu^*=3$			$\mu^*=5$		
α_1	α_1^*	T	b	$\kappa=1.2$	$\kappa=1.3$	$\kappa=1.5$	$\kappa=1.2$	$\kappa=1.3$	$\kappa=1.5$
0.1	0.5	500	80	0.191(0.182)	0.378(0.386)	0.765(0.766)	0.221(0.226)	0.394(0.402)	0.783(0.778)
		800	130	0.337(0.324)	0.622(0.633)	0.947(0.951)	0.357(0.351)	0.647(0.652)	0.974(0.978)
		1000	185	0.397(0.408)	0.714(0.709)	0.980(0.974)	0.425(0.436)	0.735(0.742)	0.998(1.000)
0.1	-0.5	500	80	0.183(0.188)	0.403(0.406)	0.763(0.775)	0.235(0.228)	0.404(0.426)	0.796(0.786)
		800	130	0.324(0.327)	0.616(0.621)	0.942(0.936)	0.343(0.338)	0.653(0.657)	0.971(0.974)
		1000	185	0.402(0.405)	0.722(0.725)	0.983(0.986)	0.418(0.427)	0.756(0.758)	1.000(0.999)

Table 3 Empirical powers of subsampling test

We now discuss the main conclusions that can be drawn from our simulation. (1) Tables 1–3 indicate that, as might be anticipated, the subsampling test produces good sizes and the powers increase as T increases. (2) The discrepancy in powers between $k^*=[0.5T]$ and $k^*=[0.75T]$ where the structural changes occur is trivial. (3) Tables 2–3 also show that if μ, α and σ change simultaneously, the powers can increase gradually as κ increases. Especially $\kappa=2$, the empirical powers can reach 1. However, the conditions can vary for μ^* and α^*. (i) In the case of a change in α, the discrepancy between $\alpha^* = 0.3$ and $\alpha^* = -0.3$ is trivial and this phenomenon also holds for $\alpha^* = \pm 0.5$. It indicates that the powers seem to be slightly less reliable on α whether or not the autoregressive parameter α is positive. The factor may be that the RCUSQ test is sensitive to σ. To illustrate, for the $\alpha = 0.3$ and $\alpha = -0.3$ case (Table 2), the subsampling tests have empirical powers of (0.241,0.237) for $T = 500$ and $\kappa = 1.2$. Whereas $\alpha = 0.3$, for the $\kappa = 1.3$ and $\kappa = 1.5$ case, the empirical powers can reach (0.448,0.883) for $T = 500$. (ii) In the case of a change in μ, the powers of the subsampling test can increase as the difference between μ and μ^* increases. In a word, the simulation evidence is strongly in favors of using our approach to the detection of changes for AR(p) parameters.

4.2 Empirical application

In this section, we analyze a group of series of daily stock price on financial assets, which contains a commercial series with SLDC (SHANGHAI LUJIAZUI FINANCE & TRADE ZONE DEVELOPMENT CO.,LTD). The stock price are observed from March 1, 2000 until June 20, 2002 with samples of $T = 540$ observations.

For tests the null of no structural changes is rejected in favour of H_1 at the 5% significance level. We consider $b = 0.15T = 81$ and find the value directly computed from the RSUCQ test statistic $\Xi_T = 3.182$ is larger than $\Xi_b^1 = 1.871$. Hence, the results give an interesting result that, for the series of SLDC, the considered period includes a structural break ($\hat{k}^* = 328$) caused by corporate restructuring on July 22, 2001 ($k^* = 330$). For these series we consider two consecutive sample of 540 observations, i.e. SLDC_1 denotes the first 328 observations for the SLDC series, while SLDC_2 denotes the subsequent 122 observation for that series. What the following figures show accords with the conclusion that the series of SLDC indeed contains structural changes in the whole samples.

The original data of SLDC_1 appears to follow the models with $Var(\varepsilon_t) = 0.27$:

$$y_t = 11.2 + \xi_t, \quad \xi_t = 0.7\xi_{t-1} + \varepsilon_t.$$

While the original data of SLDC_2 follows the other models with $Var(\varepsilon_t^*) = 1.62$:

$$y_t = 8.6 + \xi_t, \quad \xi_t = 0.9\xi_{t-1} + \varepsilon_t^*.$$

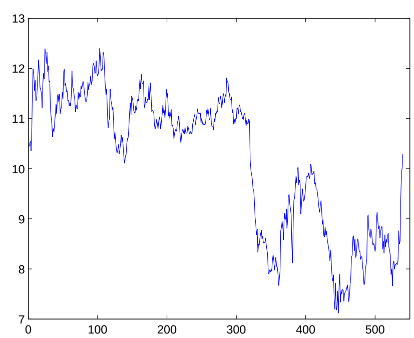

Figure 1 The original data of SLDC

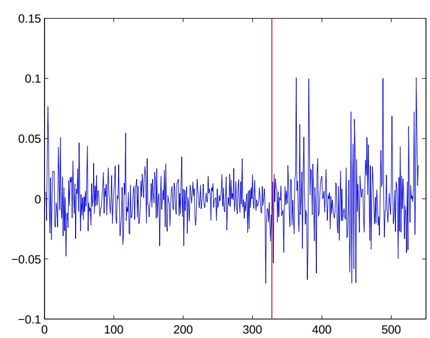

Figure 2 The first order difference data of SLDC

5. Conclusion

In this paper, the RCUSQ test for structural changes in autoregressive processes including AR(p) is proposed. We derive the asymptotic distribution of the RCUSQ test that is the function of a standard Brownian bridges. However, the test statistic is conservative to reject null hypothesis and produces low powers even for large samples. To overcome the problem, we adopt an approach based on subsampling which is a variation on the subsampling methodology of Jach and Kokoszka [2]. As most nonparametric methods, our procedure also depends on a choice of "bandwidth parameter", in our case, the subsample size b. In conclusion, the RCUSQ test based on subsampling constitutes a functional tool for detecting structural changes for AR(p) process.

Appendix A : Mathematical Proofs

Throughout the section we use the notation introduced in Sections 2 and 3. We now present a number of technical lemmas which will be needed in the proof.

Proof of Theorem 3.1 Observe that

$$\begin{aligned}\Xi_T &= \max_{q+1\le k\le T}\sqrt{T}\Big|\frac{\sum_{t=q+1}^{k}\hat{\varepsilon}_t^2}{\sum_{t=q+1}^{T}\hat{\varepsilon}_t^2}-\frac{k-q}{T-q}\Big| \\ &= \max_{0\le v\le 1}\frac{\frac{1}{\sqrt{T}}\big|\sum_{t=q+1}^{[Tv]}\hat{\varepsilon}_t^2-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}\hat{\varepsilon}_t^2\big|}{\frac{1}{T}\sum_{t=q+1}^{T}\hat{\varepsilon}_t^2}.\end{aligned}$$

Split $\hat{\varepsilon}_t^2$ into $\varepsilon_t^2+\sum_{i=1}^{5}V_{i,T}$

$$\begin{aligned}\hat{\xi}_t &= \xi_t+(\mu-\hat{\mu}) = \varepsilon_t+(\mu-\hat{\mu})+\alpha_1\xi_{t-1}+\cdots+\alpha_p\xi_{t-p}, \\ \hat{\varepsilon}_t^2 &= (\hat{\xi}_t-\hat{\alpha}_1\hat{\xi}_{t-1}-\cdots-\hat{\alpha}_p\hat{\xi}_{t-p})^2 \\ &= \varepsilon_t^2+(\mu-\hat{\mu})^2+W_T^2+2\varepsilon_t(\mu-\hat{\mu})+2W_T\varepsilon_t+2W_T(\mu-\hat{\mu}) = \varepsilon_t^2+\sum_{i=1}^{5}V_{i,T},\end{aligned}$$

where $W_T=\sum_{k=1}^{p}\alpha_k\xi_{t-k}-\sum_{k=1}^{p}\hat{\alpha}_k\hat{\xi}_{t-k}$ and $\hat{\mu}=T^{-1}\sum_{t=1}^{T}y_t$.

We first want to claim that

$$R_{i,T} =: \frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{t=q+1}^{[Tv]}V_{i,T}-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}V_{i,T}\Big| = o_p(1), \quad i=1,2,\ldots,5. \tag{2}$$

By the invariance principle for strong mixing processes [15] and Assumption 2.2, we have

$$\frac{1}{\sqrt{T}}\max_{1\le k\le T}\Big|\sum_{t=q+1}^{[Tv]}\xi_t\varepsilon_t-v\sum_{t=q+1}^{T}\xi_t\varepsilon_t\Big| = O_p(1). \tag{3}$$

First, we handle with $R_{1,T}$ and $R_{3,T}$. By Assumption 2.3

$$\hat{\xi}_t-\xi_t=\hat{\mu}-\mu=\frac{1}{T}\sum_{t=1}^{T}\xi_t=\varphi(1)\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t+o_p(1)=O_p(T^{-1/2}), \tag{4}$$

which implies that

$$\frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{t=q+1}^{[Tv]}V_{1,T}-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}V_{1,T}\Big|\le 2\frac{1}{\sqrt{T}}\Big|\sum_{t=q+1}^{T}(\mu-\hat{\mu})^2\Big|=O_p(T^{-1/2})=o_p(1)$$

and

$$\frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{t=q+1}^{[Tv]}V_{3,T}-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}V_{3,T}\Big|\le 2\frac{1}{\sqrt{T}}|\mu-\hat{\mu}|\max_{0\le v\le 1}\Big|\sum_{t=q+1}^{[Tv]}\varepsilon_t\Big|=O_p(T^{-1/2})=o_p(1).$$

Note that

$$\begin{aligned}W_T&=\sum_{k=1}^{p}\alpha_k\xi_{t-k}-\sum_{k=1}^{p}\hat{\alpha}_k\hat{\xi}_{t-k}\\&=\sum_{k=1}^{p}(\alpha_k-\hat{\alpha}_k)\xi_{t-k}+\sum_{k=1}^{p}\hat{\alpha}_k(\xi_{t-k}-\hat{\xi}_{t-k})=W_{1,T}+W_{2,T}.\end{aligned}$$

To show $R_{4,T}=o_p(1)$, it suffices to prove

$$\frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{t=q+1}^{[Tv]}\varepsilon_tW_{i,T}-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}\varepsilon_tW_{i,T}\Big|=o_p(1),\quad i=1,2. \tag{5}$$

By Lemma 2.1, (3) and (4), we have

$$\begin{aligned}&\frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{t=q+1}^{[Tv]}\varepsilon_tW_{1,T}-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}\varepsilon_tW_{1,T}\Big|\\&=\frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{k=1}^{p}(\alpha_k-\hat{\alpha}_k)\Big(\sum_{t=q+1}^{[Tv]}\varepsilon_t\xi_{t-k}-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}\varepsilon_t\xi_{t-k}\Big)\Big|\\&=O_p(T^{-1/2}\cdot p)=o_p(1),\end{aligned}$$

and

$$\begin{aligned}&\frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{t=q+1}^{[Tv]}\varepsilon_tW_{2,T}-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}\varepsilon_tW_{2,T}\Big|\\&=\frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{k=1}^{p}\hat{\alpha}_k\Big(\sum_{t=q+1}^{[Tv]}\varepsilon_t(\xi_{t-k}-\hat{\xi}_{t-k})-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}\varepsilon_t(\xi_{t-k}-\hat{\xi}_{t-k})\Big)\Big|\\&=\frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{k=1}^{p}\alpha_k\Big|+o_p(1)=O_p(T^{-1/2}\cdot p)=o_p(1).\end{aligned}$$

The proof of $R_{5,T}=o_p(1)$ is essentially the same as $R_{4,T}=o_p(1)$ and omitted for brevity.

To complete the proof, we must verify $R_{2,T}=o_p(1)$. Note $W_T=W_{1,T}+W_{2,T}$, and $W_{1,T}^2$ satisfies

$$\frac{1}{\sqrt{T}}\max_{0\le v\le 1}\Big|\sum_{t=q+1}^{[Tv]}W_{1,T}^2-\frac{[Tv]-q}{T-q}\sum_{t=q+1}^{T}W_{1,T}^2\Big|$$

$$= \frac{1}{\sqrt{T}} \max_{0 \le v \le 1} \Big| \Big(\sum_{k=1}^{p} \alpha_k - \hat{\alpha}_k \Big)^2 \Big(\sum_{t=q+1}^{[Tv]} \xi_{t-k}^2 - \frac{[Tv]-q}{T-q} \sum_{t=q+1}^{T} \xi_{t-k}^2 \Big) \Big|$$
$$= O_p(T^{-1/2}) \cdot O_p(T^{-1} \cdot p^2) \cdot O_p(T^{1/2}) = o_p(1).$$

One also can verify the negligibility of $W_{2,T}^2$ in a similar fashion to prove that of $W_{1,T}^2$. Combining these results, we have

$$\frac{1}{\sqrt{T}} \max_{0 \le v \le 1} \Big| \sum_{t=q+1}^{[Tv]} \hat{\varepsilon}_t^2 - \frac{[Tv]-q}{T-q} \sum_{t=q+1}^{T} \hat{\varepsilon}_t^2 \Big|$$
$$= \frac{1}{\sqrt{T}} \max_{0 \le v \le 1} \Big| \sum_{t=q+1}^{[Tv]} \varepsilon_t^2 - \frac{[Tv]-q}{T-q} \sum_{t=q+1}^{T} \varepsilon_t^2 \Big| + o_p(1).$$

Since $\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tv]} (\varepsilon_t^2 - E\varepsilon_t^2) \xrightarrow{P} \tau W(v)$, $\frac{v}{\sqrt{T}} \sum_{t=1}^{T} (\varepsilon_t^2 - E\varepsilon_t^2) \xrightarrow{P} v \cdot \tau W(1)$. By Assumption 2.1, and the CMT (Continuous Mapping Theorem), we can prove

$$\frac{1}{\sqrt{T}} \max_{0 \le v \le 1} \Big| \frac{\sum_{t=q+1}^{[Tv]} \varepsilon_t^2}{\sum_{t=q+1}^{T} \varepsilon_t^2} - \frac{[Tv]-q}{T-q} \Big|$$
$$= \frac{\frac{1}{\sqrt{T}} \max_{0 \le v \le 1} \Big| \sum_{t=q+1}^{[Tv]} \varepsilon_t^2 - \frac{[Tv-q]}{T-q} \sum_{t=q+1}^{T} \varepsilon_t^2 \Big|}{\frac{1}{T} \sum_{t=q+1}^{T} \varepsilon_t^2}$$
$$= \frac{\frac{1}{\sqrt{T}} \max_{0 \le v \le 1} \Big| \sum_{t=q+1}^{[Tv]} (\varepsilon_t^2 - E\varepsilon_t^2) - \frac{[Tv-q]}{T-q} \sum_{t=q+1}^{T} (\varepsilon_t^2 - E\varepsilon_t^2) \Big|}{\frac{1}{T} \sum_{t=q+1}^{T} \varepsilon_t^2}$$
$$\xrightarrow{P} \frac{\tau}{\sigma^2} \sup_{0 \le v \le 1} |BB(v)| .$$

Finally, we want to show that $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$ and $\hat{\tau}^2 \xrightarrow{P} \tau^2 = Var(\varepsilon_t^2) = E\varepsilon_t^4 - \sigma^4$. The above proofs indicate that $V_{i,T}$ satisfies

$$\frac{1}{T-q} \sum_{t=q+1}^{T} V_{i,T} = o_p(1) \quad \text{and} \quad \frac{1}{T-q} \sum_{t=q+1}^{T} V_{i,T}^2 = o_p(1). \tag{6}$$

Thus in view of $\hat{\varepsilon}_t^2 = \varepsilon_t^2 + \sum_{i=1}^{5} V_{i,T}$ and (6), we have

$$\frac{1}{T-q} \sum_{t=q+1}^{T} (\hat{\varepsilon}_t^2 - \varepsilon_t^2)^2 = \frac{1}{T-q} \sum_{t=q+1}^{T} \Big(\sum_{i=1}^{5} V_{i,T} \Big)^2 \le \frac{5}{T-q} \sum_{t=q+1}^{T} \sum_{i=1}^{5} V_{i,T}^2 = o_p(1).$$

Hence

$$\frac{1}{T} \sum_{t=1}^{T} \hat{\varepsilon}_t^2 \xrightarrow{P} E\varepsilon_t^2 = \sigma^2. \tag{7}$$

Note that

$$\frac{1}{T-q} \sum_{t=q+1}^{T} (\hat{\varepsilon}_t^2 + \varepsilon_t^2)^2 = \frac{1}{T-q} \sum_{t=q+1}^{T} (\hat{\varepsilon}_t^2 - \varepsilon_t^2)^2 + \frac{4}{T-q} \sum_{t=q+1}^{T} \varepsilon_t^2 \hat{\varepsilon}_t^2$$

$$\leq \frac{2}{T-q}\sum_{t=q+1}^{T}(\hat{\varepsilon}_t^2-\varepsilon_t^2)^2+\frac{8}{T-q}\sum_{t=q+1}^{T}\varepsilon_t^4=O_p(1)$$

and

$$\Big|\frac{1}{T-q}\sum_{t=q+1}^{T}\hat{\varepsilon}_t^4-\frac{1}{T-q}\sum_{t=q+1}^{T}\varepsilon_t^4\Big|\leq\Big(\frac{1}{T-q}\sum_{t=q+1}^{T}(\hat{\varepsilon}_t^2-\varepsilon_t^2)^2\Big)^{\frac{1}{2}}\Big(\frac{1}{T-q}\sum_{t=q+1}^{T}(\hat{\varepsilon}_t^2+\varepsilon_t^2)^2\Big)^{\frac{1}{2}}$$
$$=o_p(1)\cdot O_p(1)=o_p(1).$$

We can obtain

$$\frac{1}{T-q}\sum_{t=q+1}^{T}\hat{\varepsilon}_t^4\xrightarrow{P}\frac{1}{T-q}\sum_{t=q+1}^{T}\varepsilon_t^4=E\varepsilon_1^4.$$

This together with (6) shows that $\hat{\tau}^2\xrightarrow{P}\tau^2$. Hence, we complete the proof of the theorem. □

Before the proof of Theorem 3.2, we need introduce some important notations. Let $u_i=\varepsilon_{i+q}$, $1\leq t\leq T-q$, and

$$C_{b,l}(v)=\sum_{i=l}^{l+[(b-1)v]}u_i^2,\quad l=1,\dots,T-b-q,$$

$$C_{b,l}^1(v)=\sum_{i=l}^{l+[(b-1)v]}\varepsilon_i^{12},\quad l=1,\dots,T-b-q,$$

$$A_{b,l}=\frac{1}{b^{1/2}}\left(C_{b,l}(v)-vC_{b,l}(1)\right),\quad B_{b,l}=\frac{1}{b}C_{b,l}(1),\tag{8}$$

$$A_{b,l}^1=\frac{1}{b^{1/2}}\left(C_{b,l}^1(v)-vC_{b,l}^1(1)\right),\quad B_{b,l}^1=\frac{1}{b}C_{b,l}^1(1).\tag{9}$$

It is obvious that $\Xi_{b,l}^1=\max_{0\leq v\leq 1}\frac{|A_{b,l}^1|}{|B_{b,l}^1|}$.

Lemma A.1 *If the conditions of Theorem 3.2 hold, then*

$$\varepsilon_t^{12}-u_t^2=o_p(1),\quad t=1,\dots,T-q.$$

Proof Note that

$$\varepsilon_t^{12}-u_t^2=\varepsilon_{t+q}^{02}-\varepsilon_{t+q}^2=\hat{\varepsilon}_{t+q}^2-\varepsilon_{t+q}^2+o_p(1).$$

Since $\hat{\varepsilon}_t^2=\varepsilon_t^2+\sum_{i=1}^5 V_{i,T}$, we have

$$V_{1,T}=(\mu-\hat{\mu})^2=O_p(T^{-2})=o_p(1),$$

$$V_{3,T}=2\varepsilon_t(\hat{\mu}-\mu)=O_p(\hat{\mu}-\mu)=O_p(T^{-1/2})=o_p(1),$$

$$V_{4,T}=\varepsilon_tW_T=O_p(W_T),$$

$$V_{5,T}=W_T(\mu-\hat{\mu})=O_p(W_T)\cdot O_p(T^{-1/2}),$$

$$V_{2,T}=O_p(W_T^2).$$

It remains to prove $O_p(W_T)=o_p(1)$. By $W_T=W_{1,T}+W_{2,T}$, it is easy to show that

$$W_{1,T}=\sum_{k=1}^{p}(\alpha_k-\hat{\alpha}_k)\xi_{t-k}=O_p(T^{-1/2}\cdot p)=o_p(1),$$

$$W_{2,T}=\sum_{k=1}^{p}\hat{\alpha}_k(\xi_{t-k}-\hat{\xi}_{t-k})=\sum_{k=1}^{p}\alpha_k(\xi_{t-k}-\hat{\xi}_{t-k})+o_p(1)=O_p(T^{-1/2}\cdot p)=o_p(1).$$

Lemma A.2 *If the conditions of Theorem 3.2 hold, then*

$$\lim_{T\to\infty}\frac{1}{T-b-q}\sum_{l=1}^{T-b-q}P\{\Xi^1_{b,l}\le x\}\xrightarrow{P}G(x).$$

Proof Using the notations introduced in (8)–(9), we define

$$R^A_{b,l}=|A^1_{b,l}|-|A_{b,l}|,\quad R^B_{b,l}=|B^1_{b,l}|-|B_{b,l}|.$$

Observe that $\Xi^1_{b,l}\le x$ is equivalent to

$$|A_{b,l}|+R^A_{b,l}\le x|B_{b,l}|+xR^B_{b,l}. \tag{10}$$

Notice that (10) yields for every $\epsilon>0$ and $x>0$

$$P\{\Xi^1_{b,l}\le x\}\le P\{|A_{b,l}|\le x|B_{b,l}|+2\epsilon\}+P\{R^A_{b,l}(v)\le-\epsilon\}+P\{xR^B_{b,l}\ge\epsilon\},$$

and

$$P\{\Xi^1_{b,l}\le x\}\ge P\{|A_{b,l}|\le x|B_{b,l}|-2\epsilon\}-[P\{R^A_{b,l}(v)\ge\epsilon\}+P\{xR^B_{b,l}\le-\epsilon\}].$$

Obviously, $P\{|A_{b,l}|-x|B_{b,l}|\le 2\epsilon\}$ and $P\{|A_{b,l}|-x|B_{b,l}|\le -2\epsilon\}$ have the same limits distribution $G(x)$, as $\epsilon\to 0$. It remains to prove that for every $\epsilon>0$

$$\limsup_{T\longrightarrow\infty}\max_{1\le l\le T-b-q}P\{|R^A_{b,l}|\ge\epsilon\}=0$$

and that the same relation holds for $xR^B_{b,l}$.

We will outline the argument for $|R^A_{b,l}|\le|A^1_{b,l}-A_{b,l}|$, and

$$|A^1_{b,l}-A_{b,l}|=\frac{1}{b^{1/2}}\Big|\sum_{i=l}^{l+[(b-1)v]}(\varepsilon_i^{12}-u_i^2)-v\sum_{i=l}^{l+[(b-1)]}(\varepsilon_i^{12}-u_i^2)\Big|,$$

which is $o_p(1)$ by Lemma A.1. Hence, we can complete the proof. □

Lemma A.3 *If the conditions of Theorem 3.2 hold, then*

$$\mathrm{Var}[\tilde{G}_b(x)]\xrightarrow{P}0.$$

Proof The proof of Lemma A.3 can be immediately obtained from Jach and Kokoszka [2]. □

Proof of Theorem 3.2 To show the $\tilde{G}_b(x)\xrightarrow{p}G(x)$, it suffices to verify that $E\tilde{G}_b(x)\xrightarrow{P}G(x)$ and $Var[\tilde{G}_b(x)]\xrightarrow{p}0$. Thus Theorem 3.2 follows immediately from Lemmas A.2 and A.3. □

References

[1] BANERJEE A, LUMSDAINE R, Stock J. *Recursive and sequential tests of the unit root and trend break hypotheses: theory and international evidence* [J]. Business and Econometrica, 1994, **10**: 271–288.

[2] JACH A, KOKOSZKA P. *Subsampling unit root tests for heavy-tailed observations* [J]. Methodol. Comput. Appl. Probab., 2004, **6**(1): 73–97.

[3] PETTITT A N. *A nonparametric approach to the change-point problem* [J]. J. Roy. Statist. Soc. Ser. C, 1979, **28**(2): 126–135.

[4] SEN A, SRIVASTAVA M S. *On tests for detecting change in mean* [J]. Ann. Statist., 1975, **3**: 98–108.

[5] PICARD D. *Testing and estimating change-points in time series* [J]. Adv. in Appl. Probab., 1985, **17**(4): 841–867.

[6] POLITIS D, ROMANO J P, WOLF M. *Subsampling* [M]. New York: Springer, 1999.

[7] ANDREWS D W K. *Tests for parameter instability and structural change with unknown change point* [J]. Econometrica, 1997, **61**: 821–856.

[8] GOMBAY E. *Change detection in autoregressive time series* [J]. J. Multivariate Anal., 2008, **99**(3): 451–464.

[9] PAGE E S. *A test for change in a parameter occurring at an unknown point* [J]. Biometrika, 1955, **42**: 523–527.

[10] CHOW G C. *Test of equality between sets of coefficients in two linear regressions* [J]. Econometrica, 1960, **28**: 591–605.

[11] CHERNOFF H, ZACKS S. *Estimating the current mean of a normal distriubtion which is subjected to changes in time* [J]. Ann. Math. Statist., 1964, **35**: 999–1018.

[12] BAI J. *Estimation of a change point in multiple regression models* [J]. Time Series Analysis, 1997, **15**: 453–472.

[13] GIRAITIS L, LEIPUS R, SURGAILIS D. *The change-point problem for dependent observations* [J]. J. Statist. Plann. Inference, 1996, **53**(3): 297–310.

[14] CSÖRGŐ M, HORVÁTH L. *Limits Theorems in Change-Point Analysis* [M]. New York, Wiley, 1997.

[15] PLIGRAD M. *Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables* [J]. Birkhäuser Boston, Boston, MA, 1986.

[16] KOKOSZKA P, LEIPUS R. *Change-point in the mean of dependent observations* [J]. Statist. Probab. Lett., 1998, **40**(4): 385–393.

[17] DAVIS R, RESNICK S. *Limit theory for the sample covariance and correlation functions of moving averages* [J]. Ann. Statist., 1986, **14**(2): 533–558.

[18] DAVIS R A, HUANG Dawei, YAO Yiching. *Testing for a change in the parameter values and order of an autoregressive model* [J]. Ann. Statist., 1995, **23**(1): 282–304.

[19] DAVIS R A, HSING T. *Point process and partial sum convergence for weakly dependent random variables with infinite variance* [J]. Ann. Probab., 1995, **23**(2): 879–917.

[20] BROWN R L, DURBIN J, EVANS J M. *Techniques for testing the constancy of regression relationships over time* [J]. J. Roy. Statist. Soc. Ser. B, 1975, **37**: 149–192.

[21] BEVERIDGE S, NELSON C R. *A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of 'Business Cycle'* [J]. Journal of Monetary Economics, 1981, **7**: 151–74.

[22] LEE S, PARK S. *The cusum of squares test of test for scale changes in infinite order moving average processes* [J]. Scandinavian Journal of Statistics, 2001, **28**: 625–644.

[23] YOHAI V J, MARONNA R A. *Asymptotic behavior of least-squares estimates for autoregressive processes with infinite variances* [J]. Ann. Statist., 1977, **5**(3): 554–560.

[24] VOGELSANG T J. *Source of nonmonotonic power when testing for a shift in mean of a dynamic time series* [J]. Econometrics, 1999, **88**: 283–299.