Journal of Mathematical Research & Exposition
Jul., 2010, Vol. 30, No. 4, pp. 610–618
DOI:10.3770/j.issn:1000-341X.2010.04.005
Http://jmre.dlut.edu.cn

Sequential Monitoring Variance Change in Linear Regression Model

Zhan Shou CHEN[1,*], Zheng TIAN[1,2], Rui Bing QIN[1], Cheng Cai LENG[1]
1. *Department of Applied Mathematics, Northwestern Polytechnical University, Shaanxi* 710129*, P. R. China;*
2. *State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing* 100101*, P. R. China*

Abstract The paper investigates the sequential observations' variance change in linear regression model. The procedure is based on a detection function constructed by residual squares of CUSUM and a boundary function which is designed so that the test has a small probability of false alarm and asymptotic power one. Simulation results show our monitoring procedure performs well when variance change occurs shortly after the monitoring time. The method is still feasible for regression coefficients change or both variance and regression coefficients change problem.

Keywords sequential monitoring; variance change; linear regression model; residuals.

Document code A
MR(2000) Subject Classification 62F12; 60P20
Chinese Library Classification O212.1

1. Introduction

The problem of testing for a parameter change has attracted much attention from many researchers since the parameter change in the underlying model is occasionally observed in actual practice. Change-point detection procedures fall into two categories: retrospective or a posteriori tests and on-line, sequential or priori tests. The former tests study a fixed historical sample. The latter tests, which our paper will use, is motivated by the work of Chu, Stinchcombe and White [1] who proposed the following problem: Given a previously estimated model, the arrival of new data invites the question: is yesterday's model capable of explaining today's data? By using this idea, Berkes et al. [2] detected change-point in GARCH(p, q) models, Aue [3, 4] tested changes in RCA(1) time series, Zeileis et al. [5] studied structural change in dynamic econometric models, and Andreous and Ghysels [6] detected disruptions in financial markets.

The problem of testing for a variance change has become an important issue in time series analysis since the variance is often interpreted as a risk in econometrics. There are many papers considering this problem, for example, Inclán and Tiao [7], Csörgo and Horváth [8], Lee and

Received October 9, 2008; Accepted September 15, 2009
Supported by the National Natural Science Foundation of China (Grant Nos. 60972150; 10926197) and the Science and Technology Innovation Foundation of Northwestern Polytechnical University (Grant No. 2007KJ01033).
* Corresponding author
E-mail address: chenzhanshou@126.com (Z. S. CHEN)

Na [9], among many others. However, all these papers test for fixed historical samples. About sequential tests of variance change problem, Carsoule et al. [10] detected normal random variable data, Horváth et al. [11] investigated on-line detection of change in unconditional variance in a conditionally heteroskedastic time series. The change point problem of linear regression as a most useful and simple model was studied by many authors. For instance, Bai [12], Lee et al. [13]. Horváth et al. [14, 15], Aue et al. [16] monitored change point of regression parameter in linear model. However, few literature monitored variance change in linear regression model. Motivated by this, in this paper we discuss sequential variance change monitoring problem in linear regression model.

Our monitoring procedure is based on a detection function constructed by residual squares of CUSUM and a boundary function which is designed so that the test has a small probability of false alarm and asymptotic power one. Simulation results show that our monitoring procedure performs well when change occurs shortly after the monitoring time. Such a formulation is, for example, relevant to that there is a need to quickly detect a change in the variance of returns which corresponds to a change in the behavior of the fluctuation of stock price or trading volume from small to big, or vice versa. We also test our procedure's performance for monitoring regression coefficients change or both regression coefficients and variance change problem. The simulation results show that our method is also sensitive for these change points.

The paper is organized as follows. In Section 2, we introduce our model assumptions and sequential testing approach. The main results of the paper and some discussion about boundary are given in Section 3. Section 4 reports simulation results. A simple conclusion is given in Section 5 and all proofs of main results are gathered into Section 6.

2. Monitoring process and assumptions

We consider the linear regression model $y_i = \mathbf{x}_i^{\mathrm{T}}\beta_i + \varepsilon_i,\ 1 \le i < \infty$, where $\mathbf{x}_i$ is a $p \times 1$ dimensional i.i.d. random vector, β_i is a $p \times 1$ dimensional random parameter vector and $\{\varepsilon_i\}$ is an i.i.d. error sequence with

$$E\varepsilon_i = 0,\ E\varepsilon_1^2 = \sigma_0^2 < \infty,\ \text{and}\ \mathrm{Var}(\varepsilon_i) = \sigma_i^2 < \infty. \tag{1}$$

Since we are interested in monitoring variance change, we assume regression parameter as nuisance parameter and will not change, namely, $\beta_i = \beta_0,\ 1 \le i < \infty$.

First we give a "non-contamination assumption", namely, there is no change in variance during the first m observations, i.e.,

$$\sigma_1^2 = \cdots = \sigma_m^2 \equiv \sigma_0^2. \tag{2}$$

Next, observing new data, we want to detect if a change occurs in the variance. That is, we want to test the null hypothesis

$$H_0:\ \sigma_i^2 = \sigma_0^2, \quad i = m+1, m+2, \ldots, \tag{3}$$

against the alternative

$$H_A: \text{there is a } k^* \ge 1 \text{ such that } \sigma_i^2 = \sigma_0^2,\ i = m+1, m+2, \ldots, m+k^*,$$

$$\text{but } \sigma_i^2 = \sigma_A^2,\ i = m + k^* + 1, m + k^* + 2, \ldots,\ \text{with } \sigma_A^2 \neq \sigma_0^2. \tag{4}$$

The parameters σ_0^2, σ_A^2 and k^*, the so-called change-point, are assumed unknown.

The monitoring procedure presented in this paper uses a detector function $Q(m,k)$ and a boundary function $g(m,k)$ which together define the stopping time

$$\tau(m) = \inf\{k \geq 1,\ |Q(m,k)| \geq cg(m,k)\}$$

(with the understanding that $\inf \phi = \infty$) which must satisfy

$$\begin{aligned} &\lim_{m\to\infty} P\{\tau(m) < \infty\} = \alpha,\ \text{under } H_0; \\ &\lim_{m\to\infty} P\{\tau(m) < \infty\} = 1,\ \text{under } H_A. \end{aligned} \tag{5}$$

The index k labels the time elapsed after the monitoring has commenced. The probability $\alpha \in (0,1)$ controls the false alarm rate. In order to give our results, we need more assumptions on the regression model:

$$\{\varepsilon_i, 1 \leq i < \infty\} \text{ and } \{\mathbf{x}_i, 1 \leq i < \infty\} \text{ are independent.} \tag{6}$$

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^{\mathrm{T}} \longrightarrow \mathsf{C} > 0,\ \text{a.s.} \tag{7}$$

3. Main results

From initial observations $\mathbf{x}_1, \ldots, \mathbf{x}_m$, let

$$\hat{\beta}_m = (\sum_{i=1}^{m} \mathbf{x}_i \mathbf{x}_i^{\mathrm{T}})^{-1} \sum_{j=1}^{m} \mathbf{x}_j y_j$$

be the ordinary least squares estimator at time m. Then the resduials are

$$\hat{\varepsilon}_i = y_i - \mathbf{x}_i^{\mathrm{T}} \hat{\beta}_m, \quad 1 \leq i < \infty. \tag{8}$$

We use the CUSUM of residual squares as the monitoring function

$$Q(m,k) = \sum_{i=m+1}^{m+k} (\hat{\varepsilon}_i^2 - \hat{\theta}_m^2)$$

and

$$g(m,k) = m^{1/2}(1 + \frac{k}{m})(\frac{k}{m+k})^{\gamma}, \quad 0 \leq \gamma < \frac{1}{2}$$

as our boundary function, in which $\hat{\theta}_m^2 = m^{-1}\sum_{i=1}^{m} \hat{\varepsilon}_i^2$. Then we have following theorems.

Theorem 1 *Assume (1), (6) and (7) hold. Then under the null hypothesis (i.e., (2)) we have*

$$\lim_{m\to\infty} P\Big\{\sup_{1\leq k<\infty} \frac{|Q(m,k)|}{\sigma g(m,k)} \leq c\Big\} = P\Big\{\sup_{0\leq t\leq 1} \frac{|W(t)|}{t^{\gamma}} \leq c\Big\}, \tag{9}$$

where $\{W(t),\ 0 \leq t < \infty\}$ denotes a Wiener process, and $c = c(\alpha)$.

Theorem 2 *Assume (1), (6) and (7) hold. Then under the alternative hypothesis (i.e., (3)) we* *have*

$$\lim_{m\to\infty}\sup_{1\le k<\infty}\frac{|Q(m,k)|}{\sigma g(m,k)}\xrightarrow{p}\infty. \tag{10}$$

Since σ is unknown in practice, we could replace it by estimator

$$\hat{\sigma}^2=m^{-1}\sum_{i=1}^{m}\hat{\varepsilon}_i^4-(m^{-1}\sum_{i=1}^{m}\hat{\varepsilon}_i^2)^2.$$

One can easily prove that

$$\hat{\sigma}\xrightarrow{p}\sigma. \tag{11}$$

Then from the Slutsky theorem, the assertions of Theorems 1 and 2 still hold when σ is replaced by $\hat{\sigma}$.

The choice of the boundary function in sequential test is traditionally based on the minimization of the average detection delay of the average run length (ARL). The paper's boundary $b(t)=(1+t)(t/(1+t))^\gamma$ is similar to Horváth et al. [14] and other related papers'. Chu et al. [1] and Carsoule [10] used $b_1(t)=\sqrt{t(a^2+\log t)}$ as their boundary, where $t\ge 1$ and $a^2=7.78$ and 6.25 for nominal level $\alpha=0.05$, 0.1, respectively. For dynamic econometric model, Zeileis et al. [5] suggested boundary $b_2(t)=\lambda\sqrt{\log_+(1+t)}$, where λ relies on the monitoring horizon. Although the papers' asymptotic distribution itself is not a Wiener process, in fact according to the conclusion

$$\{W^0(t),\ t>1\}\overset{d}{=}\{(t-1)W(\frac{t}{t-1}),\ t>1\},$$

and a simple computation [10], we also could use Chu's [1] boundary as our monitoring boundary. Since boundary $b_2(t)$ is defined for "closed end" stopping rules, and our stopping rule is an "open ended" problem, we use finite sample simulation to compare monitoring performance just only for boundary $b(t)$ and $b_1(t)$.

4. Simulations

In this section, we illustrate the theory developed in the previous section to assess the finite sample performance of our monitoring schemes by 2500 replications. We consider the model $y_i=\beta x_i+\varepsilon_i$, and generate innovation data from i.i.d.$N(0,1)$ random variables. We compute the empirical crossing probabilities under H_0 for historical sample sizes $m=50$, 100, 200 and 300. The monitoring horizon q is set to be two, four and eight times the historical sample size. Table 1 reports the empirical sizes of four different boundaries for nominal level $\alpha=0.05$, 0.1, respectively. Where b_1, b_2 and b_3 denote to choose $\gamma=0,\ 0.25,\ 0.49$ in our boundary respectively, and b_4 denotes to use Chu's [1] boundary $b_1(t)$.

From Table 1 one can easily find, in all cases, except for the 5 percent test with $\gamma=0.49$, the empirical sizes do not exceed the nominal sizes as predicted by the theory in the previous sections. Furthermore the empirical sizes even for big monitoring horizon are very small for relative smaller γ. That is to say, if the monitoring horizon takes place over interval of fixed length, then the probability of false rejection is very small and reduces while increasing historical size m or choosing γ close to zero.

		$q = 2m$				$q = 4m$				$q = 8m$			
α	m	b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
5%	50	5.09	0.95	0.01	0.1	5.74	1.65	0.02	0.07	5.75	1.9	0.05	0.225
	100	4.31	0.22	0	0.01	5.22	0.6	0.07	0.15	5.38	0.9	0.09	0.86
	200	4.7	0.38	0	0	5.59	0.46	0.01	0.06	5.17	0.62	0.04	0.16
	300	5.9	0.8	0	0	5.92	0.18	0	0.02	5.26	0.39	0	0.17
10%	50	6.87	1.96	0.2	0.19	7.12	2.88	0.8	1.38	7.63	3.66	1.23	3.55
	100	6.68	1.2	0.04	0.08	6.55	1.1	0.03	0.15	7.56	3.05	0.69	1.38
	200	7.47	0.82	0.04	0.05	7.23	0.89	0	0.06	9.56	1.52	0.34	0.65
	300	7.49	0.61	0	0	7.22	0.74	0.04	0.08	7.55	1.55	0.11	0.31

Table 1 Empirical size

To study the finite sample power of the test procedure, we simulate a variance change from 1 to 2 at $k^* = 3$, (since our monitoring procedure is symmetric, if there is a decrease change in variance, there will be similar test power and delay time for same change time and change size. We omit the simulation results here). When there is a change, we need to quickly give an alarm (small average run length), so we set monitoring horizon q to be $\frac{1}{4}$, $\frac{1}{2}$ and one times the historical sample size. According to Table 2 we could easily find, when shift occurs at the beginning time of monitoring procedure, it will be more sensitive to choose relatively bigger γ, and bigger historical sample size gives higher asymptotic powers. It is clear to see that our boundary always performs better than Chu's in asymptotic power and delay time.

		$q = \frac{1}{4}m$				$q = \frac{1}{2}m$				$q = m$			
α	m	b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
5%	50	2.7	0.03	0	0	21.7	9.7	0.12	0	49.7	45.6	28.3	27.4
	100	21.9	3.82	0	0	53.2	40.7	12.9	4.72	73.5	67.2	53.2	50
	200	55.3	33.9	0.21	0	74.4	63.2	42.2	34.8	86.1	80.4	69.6	65.9
	300	67.7	48.1	12.2	0.6	82.3	72.2	53.7	47.1	91.4	85.9	76.6	73.1
10%	50	4.34	0.72	0	0	26.1	19.1	1	0.46	54.5	52.6	38.6	34.4
	100	30.1	10.7	0	0	58.6	49.7	26.7	15.8	75.8	71.7	59.6	54.9
	200	59.1	42.2	5.27	0	77	68.4	50.1	41.2	87.8	83.6	74.6	69.6
	300	70.2	55.2	24.9	9.76	84.5	76.6	60.4	52.7	91.9	88	80.2	76.3

Table 2 Asymptotic power when variance changes from 1 to 2 at $k^* = 3$

Since regression coefficients, variance or both of these parameters in the model may change in practical problems, we now test procedure's performance for monitoring these changes by simulation. To save the space, we just only report some results for nominal level $\alpha = 0.1$. Table 3 gives the asymptotic power when regression coefficient β changes from 2 to 2.8 at $k^* = 3$. The results show the procedure can also be used to monitor change point of regression coefficients, but the alarm time will be longer, namely, it has bigger average run length. Table 4 presents the simulation results when both variance changes from 1 to 2 and coefficient changes from 2 to 2.8 at $k^* = 3$. In this case the procedure can get a similar asymptotic power by a smaller monitoring delay time.

	$q=\frac{1}{2}m$				$q=m$				$q=2m$			
m	b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
50	22.1	14.9	5.3	3.1	35.9	31.3	18.6	15.7	49.6	49.3	39.6	41.6
100	35.9	25.3	10.1	6.35	53.3	48.1	32.7	28.7	69.8	69.6	59.6	59.9
200	56.9	42.7	21.8	14.6	74.8	68.5	53.8	48.9	86.8	84.3	76.3	74.1
300	69.2	56.4	33.7	25.1	83.9	77.8	65.4	60.3	91.5	88.4	82.1	79.6

Table 3 Asymptotic power when regression coefficient changes from 2 to 2.8 at k^*

	$q=\frac{1}{2}m$				$q=\frac{3}{4}m$				$q=m$			
m	b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
50	52.6	46.8	29.5	22.8	67.3	63.9	51.4	45.7	74.7	72.1	62.1	57.8
100	75.6	69.2	53.9	46.5	82.9	78.7	68.5	63.6	86.8	83.7	75.7	72.2
200	87.3	81.7	69.7	63.9	91.5	87.8	79.4	75.3	93.3	90.4	84.3	81.3
300	91.3	86.3	75.8	70.9	94.1	90.8	83.9	80.4	95.6	93.1	87.8	85.2

Table 4 Asymptotic power when both coefficient and variance change at $k^* = 3$

In summary, if a process is to be monitored for long time and type I error is to be avoided, we could choose γ close to 0. On the other hand, if it is important to detect a change as soon as possible and if this change can be expected to occur shortly after the beginning time of monitoring and some false alarms is acceptable, then it is better to choose γ close to 0.5, and for all of these case, a bigger history samples give a higher power. The procedure can also be used to monitor regression coefficients change point or all of them, but delay time when there just occurs coefficients change will be longer than when there only occurs variance change.

5. Conclusions

In this paper we proposed a sequential monitoring procedure to examine a structural break in the variance of linear regression model with i.i.d. innovations. We showed through simulations that our procedure has good empirical size, power properties and relatively small delay time when change occurs at the early stage of monitoring and is still feasible when regression coefficients occur break or both regression coefficients and variance change.

Testing for breaks is closely connected with model specification analysis as the forecasting. Out-of-sample prediction is typically based on the maintained assumption of model (parameter) stability. Sequential analysis is therefore a desirable tool when out-of-sample analysis is performed since it allows for real-time monitoring of prediction models. Moreover, monitoring financial risk can be considered as a useful statistical result showing our method is useful. However, defining a boundary powerful for all monitoring time is still a hard job. Delay time of monitoring and distribution of stopping times for this problem is meaningful and we will study it in latter articles.

6. Proof of main results

Proof of Theorem 1 Since $\hat{\beta}_m$ is the OLS of β_0, we have that

$$\sqrt{m}(\hat{\beta}_m - \beta_0) = O_p(1). \tag{12}$$

From (8) we could easily find that

$$\begin{aligned}
Q(m,k) = & \sum_{i=m+1}^{m+k} (\hat{\varepsilon}_i^2 - \hat{\theta}_m^2) = \sum_{i=m+1}^{m+k} [\varepsilon_i - \mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m - \beta_0)]^2 - \frac{k}{m}\sum_{i=1}^{m}[\varepsilon_i - \mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m - \beta_0)]^2 \\
= & \Big(\sum_{i=m+1}^{m+k} \varepsilon_i^2 - \frac{k}{m}\sum_{i=1}^{m}\varepsilon_i^2\Big) + \Big(\sum_{i=m+1}^{m+k}[\mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m - \beta_0)]^2 - \frac{k}{m}\sum_{i=1}^{m}[\mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m - \beta_0)]^2\Big) - \\
& 2\Big(\sum_{i=m+1}^{m+k}(\hat{\beta}_m - \beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i - \frac{k}{m}\sum_{i=1}^{m}(\hat{\beta}_m - \beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i\Big).
\end{aligned}$$

Then

$$\sup_{1\le k<\infty} |Q(m,k)|/\{\sigma g(m,k)\} \le I_1 + I_2 + 2I_3, \tag{13}$$

where

$$\begin{aligned}
I_1 &= \sup_{1\le k<\infty}\Big|\sum_{i=m+1}^{m+k}\varepsilon_i^2 - \frac{k}{m}\sum_{i=1}^{m}\varepsilon_i^2\Big|/\{\sigma g(m,k)\}, \\
I_2 &= \sup_{1\le k<\infty}\Big|\sum_{i=m+1}^{m+k}[\mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m - \beta_0)]^2 - \frac{k}{m}\sum_{i=1}^{m}[\mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m - \beta_0)]^2\Big|/\{\sigma g(m,k)\}, \\
I_3 &= \sup_{1\le k<\infty}\Big|\sum_{i=m+1}^{m+k}(\hat{\beta}_m - \beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i - \frac{k}{m}\sum_{i=1}^{m}(\hat{\beta}_m - \beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i\Big|/\{\sigma g(m,k)\}.
\end{aligned}$$

By (7) and (12),

$$\begin{aligned}
I_2 &\le \sup_{1\le k<\infty}\Big(\sum_{i=m+1}^{m+k}||\mathbf{x}_i||^2||\hat{\beta}_m - \beta_0||^2 + \frac{k}{m}\sum_{i=1}^{m}||\mathbf{x}_i||^2||\hat{\beta}_m - \beta_0||^2\Big)^2/\{\sigma g(m,k)\} \\
&= \sup_{1\le k<\infty}\Big(\sum_{i=1}^{m+k}||\mathbf{x}_i||^2||\hat{\beta}_m - \beta_0||^2 + \frac{m+k}{m}\sum_{i=1}^{m}||\mathbf{x}_i||^2||\hat{\beta}_m - \beta_0||^2\Big)^2/\{\sigma g(m,k)\} \\
&= O_p(1)\sup_{1\le k<\infty}\frac{2(m+k)/m}{m^{1/2}(1+\frac{k}{m})(\frac{k}{m+k})^{\gamma}} = o_p(1).
\end{aligned} \tag{14}$$

By the central limit theorem and (6), we get as $m\to\infty$ that

$$\sum_{i=1}^{m}\mathbf{x}_i\varepsilon_i = O_p(m^{1/2}). \tag{15}$$

Then according to (12), (15) and proposition 6.1.1 of Brockwell and Davis [17] we have

$$\begin{aligned} I_3 &= \sup_{1\le k<\infty} \Big|\sum_{i=1}^{m+k}(\hat{\beta}_m-\beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i - \frac{m+k}{m}\sum_{i=1}^{m}(\hat{\beta}_m-\beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i\Big|/\{\sigma g(m,k)\} \\ &\le \sup_{1\le k<\infty}\Big(\Big|\sum_{i=1}^{m+k}(\hat{\beta}_m-\beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i\Big| - \frac{m+k}{m}\Big|\sum_{i=1}^{m}(\hat{\beta}_m-\beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i\Big|\Big)/\{\sigma g(m,k)\} \\ &= O_p(1)\sup_{1\le k<\infty}\frac{\sqrt{(m+k)/m}+(m+k)/m}{m^{1/2}(1+\frac{k}{m})(\frac{k}{m+k})^{\gamma}} = o_p(1). \end{aligned} \tag{16}$$

By (1) we have

$$\{\sum_{i=m+1}^{m+k}(\varepsilon_i^2-\sigma_0^2),\ 1\le k<\infty\} \text{ and } \{\sum_{i=1}^{m}(\varepsilon_i^2-\sigma_0^2)\} \text{ are indenpendent for each } m.$$

Hence by the functional center limit theorem we can find two independent Wiener processes $\{W_{1,m}(t)\}$ and $\{W_{2,m}(t)\}$ such that

$$m^{-1/2}\sum_{i=m+1}^{m+k}(\varepsilon_i^2-\sigma_0^2) \xrightarrow{d} \sigma W_{1,m}(\frac{k}{m}),\quad m\to\infty,$$

and

$$m^{-1/2}\sum_{i=1}^{m}(\varepsilon_i^2-\sigma_0^2) \xrightarrow{d} \sigma W_{2,m}(1),\quad m\to\infty.$$

Then

$$\sum_{i=m+1}^{m+k}\varepsilon_i^2 - \frac{k}{m}\sum_{i=1}^{m}\varepsilon_i^2 \xrightarrow{d} m^{1/2}\sigma\Big(W_{1,m}(\frac{k}{m}) - \frac{k}{m}W_{2,k}(1)\Big),$$

which leads to

$$I_1 \xrightarrow{d} \sup_{1\le t<\infty}\frac{|W_1(t)-tW_2(1)|}{(1+t)(\frac{t}{1+t})^{\gamma}}.$$

Berkes et al. [2] showed that for all m,

$$\sup_{1\le t\le\infty}\frac{|W_1(t)-tW_2(1)|}{(1+t)(\frac{t}{1+t})^{\gamma}} \xrightarrow{d} \sup_{1\le t\le 1}\frac{|W(t)|}{t^{\gamma}}. \tag{17}$$

Therefore, Theorem 1 follows immediately from (13), (14), (16) and (17). □

Proof of Theorem 2 Let $\tilde{k}=k^*+m$. Then

$$\begin{aligned} \frac{Q(m,\tilde{k})}{m} =& \frac{1}{m}\Big(\sum_{i=m+1}^{m+\tilde{k}}\varepsilon_i^2 - \frac{\tilde{k}}{m}\sum_{i=1}^{m}\varepsilon_i^2\Big) - \frac{2}{m}\Big(\sum_{i=m+1}^{m+\tilde{k}}(\hat{\beta}_m-\beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i - \frac{\tilde{k}}{m}\sum_{i=1}^{m}(\hat{\beta}_m-\beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i\Big) + \\ & \frac{1}{m}\Big(\sum_{i=m+1}^{m+\tilde{k}}[\mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m-\beta_0)]^2 - \frac{\tilde{k}}{m}\sum_{i=1}^{m}[\mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m-\beta_0)]^2\Big) = I_4+I_5+I_6. \end{aligned} \tag{18}$$

Since regression parameter as a nuisance parameter will not change, from the proof of Theorem 1 we have

$$I_5 = -\frac{2}{m}\Big(\sum_{i=m+1}^{m+\tilde{k}}(\hat{\beta}_m-\beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i - \frac{\tilde{k}}{m}\sum_{i=1}^{m}(\hat{\beta}_m-\beta_0)^{\mathrm{T}}\mathbf{x}_i\varepsilon_i\Big) = o_p(1), \tag{19}$$

$$I_6 = \frac{1}{m}\Big(\sum_{i=m+1}^{m+\tilde{k}} \big[\mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m - \beta_0)\big]^2 - \frac{\tilde{k}}{m}\sum_{i=1}^{m} \big[\mathbf{x}_i^{\mathrm{T}}(\hat{\beta}_m - \beta_0)\big]^2\Big) = o_p(1). \tag{20}$$

By (1), (4) and the invariance principle, we have

$$\begin{aligned} I_4 &= m^{-1}\Big(\sum_{i=m+1}^{m+\tilde{k}} \varepsilon_i^2 - \frac{\tilde{k}}{m}\sum_{i=1}^{m}\varepsilon_i^2\Big) \\ &= m^{-1}\Big(\sum_{i=m+1}^{m+k^*} (\varepsilon_i^2 - \sigma_0^2) + \sum_{i=m+k^*+1}^{m+\tilde{k}} (\varepsilon_i^2 - \sigma_A^2) - \frac{\tilde{k}}{m}\sum_{i=1}^{m}(\varepsilon_i^2 - \sigma_0^2)\Big) + (\sigma_A^2 - \sigma_0^2) \\ &= (\sigma_A^2 - \sigma_0^2) + o_p(1). \end{aligned} \tag{21}$$

Combining (18)–(21), we conclude

$$\liminf_{m\to\infty} \frac{|Q(m,\tilde{k})|}{m(1+k/m)(k/(m+k))^{\gamma}} > 0.$$

This completes the proof of Theorem 2. □

References

[1] CHU C S J, STINCHCOMBE M, WHITE H. *Monitoring structural change* [J]. Econometric, 1996, **64**: 1045–1065.

[2] BERKES I, GOMBAY E, HORVÁTH L. et al. *Sequential change-point detection in* GARCH(p, q) *models* [J]. Econometric Theory, 2004, **20**(6): 1140–1167.

[3] AUE A. *Strong approximation for RCA(1) time series with applications* [J]. Statist. Probab. Lett., 2004, **68**(4): 369–382.

[4] AUE A. *Testing for parameter stability in RCA(1) time series* [J]. J. Statist. Plann. Inference, 2006, **136**(9): 3070–3089.

[5] ZEILEIS A, LEISCH F, KLEIBER C. et al. *Monitoring structural change in dynamic econometric models* [J]. J. Appl. Econometrics, 2005, **20**(1): 99–121.

[6] ANDREOU E, GHYSELS E. *Monitoring disruptions in financial markets* [J]. J. Econometrics, 2006, **135**(1-2): 77–124.

[7] INCLÁN C, TIAO G C. *Use of cumulative sums of squares for retrospective detection of changes of variance* [J]. J. Amer. Statist. Assoc., 1994, **89**(427): 913–923.

[8] CSÖRGO M, HORVÁTH L. *Limit Theorems in Change-Point Analysis* [M]. Wiley, New York, 1997.

[9] LEE S, NA O, NA S. *On the Cusum squares test for variance change in nonstationary and nonparametric time series models* [J]. Ann. Inst. Statist. Math., 2003, **55**(3): 467–485.

[10] CARSOULE F, FRANSES P H. *Monitoring structural change in variance with an application to European nominal exchange rate volatility* [C]. Econometric Institute Report, 1995, **154**.

[11] HORVÁTH L, KOKOSZKA P, ZHANG Aonan. *Monitoring constancy of variance in conditionally heteroskedastic time series* [J]. Econometric Theory, 2006, **22**(3): 373–402.

[12] BAI Jushan. *Weak convergence of the sequential empirical processes of residuals in ARMA models* [J]. Ann. Statist., 1994, **22**(4): 2051–2061.

[13] LEE S, TOKUTSU Y, MAEKAWA K. *The cusum test for parameter change in regression models with ARCH errors* [J]. J. Japan Statist. Soc., 2004, **34**(2): 173–188.

[14] HOVÁTH L, HUŠKOVÁ M, KOKOSZKA P. *Monitoring changes in linear models* [J]. J. Statist. Plann. Inference, 2004, **126**(1): 225–251.

[15] HORVÁTH L, KOKOSZKA P, STEINEBACH J. *On sequential detection of parameter changes in linear regression* [J]. Statist. Probab. Lett., 2007, **77**(9): 885–895.

[16] AUE A, HORVÁTH L, HUŠKOVÁ M. et al. *Change-point monitoring in linear models* [J]. Econom. J., 2006, **9**(3): 373–403.

[17] BROCKWELL P J, DAVIS R A. *Time Series: Theory and Methods* [M]. Spring, New York, 1991.