Hindawi
Journal of Mathematics
Volume 2021, Article ID 1455812, 6 pages
https://doi.org/10.1155/2021/1455812

Research Article

A Two-Stage Estimator for Change Point in the Mean of Panel Data

Wenzhi Zhao, Yinqian Yang, and Di Zhang

School of Science, Xi'an Polytechnic University, Xi'an, Shaanxi, China

Correspondence should be addressed to Wenzhi Zhao; zhaowenzhi1979@126.com

Received 22 August 2021; Accepted 2 September 2021; Published 18 September 2021

Academic Editor: Wenpeng Zhang

In this paper, a two-stage consistency estimator for change point in the mean of panel data is given. Firstly, a single sequence is extracted, and the initial estimator and confidence interval of the change point are given by the least square method. Based on the confidence interval, a random interval containing change point with probability tending to 1 is constructed. Secondly, using all panel data falling into the random interval, the final estimator of change point is obtained by least square estimation. The asymptotic distribution is established. Simulation results show that our method can not only ensure the estimation accuracy but also greatly reduce time complexity.

1. Introduction

One of the tools to analyse large, high-dimensional datasets is the panel data model. This paper studies the problem of structural changes for panel data, in which there are N series (variables), and each series has T observations. It is assumed that there is only one change that has taken place in each series at an unknown common point, referred to as the common change point. Common change points in panel data are wide spread phenomena. For example, the outbreak of the epidemic may impact every country's GDP. A tax policy change may alter each firm's investment incentive. While it may be difficult to identify a change point with single series, it should be, naturally, much easier to locate the common change point using a number of series together. This paper explores the panel data approach to the estimation of change point.

Joseph and Wolfson [1, 2] are the early researchers who laid the groundwork in change point for panel data. They proposed a random change point in which each series has its own change point; across N series, the change points are assumed to be independent and identically distributed (i.i.d.). They proved that the common distribution of the i.i.d. change points can be consistently estimated. This random change point model is extended to the autoregressive model proposed by Joseph et al. [3]. Joseph et al. [4] considered the Bayesian framework. Skates et al. [5] and Jackson and Sharples [6] studied the application-oriented Bayesian models. Bai [7] established the consistency of the estimated common change point in panel data by the least square method. Horváth and Hušková [8] and Shin and Hwang [9] used the CUSUM method to test change point in the mean of panel data. Bai [7] conducted ratio-type statistics to detect change point in panel data.

Computer-based technology allows scientists to collect enormous datasets, and huge data demand new methodology. In massive data, how to quickly and accurately estimate the change point location has become a real problem to be solved. Cao and Xia [10] considered a fast estimation method for univariate sequence. This paper proposes a two-stage estimation method to locate change point in the mean of panel data, and the consistency of the estimator is proved.

2. Model and Assumptions

We assume that we study N panels and we have T observations in each panel. We define our model as

$$X_{it}=\begin{cases}\mu_{i1}+\varepsilon_{it}, & 1\le t\le k_0,\\ \mu_{i2}+\varepsilon_{it}, & k_0+1\le t\le T,\end{cases}\quad t=1,2,\ldots,T,\ i=1,2,\ldots,N, \tag{1}$$

where $E\varepsilon_{it}=0$ for all i and t. In this model, each series has a change point at k_0, where k_0 is unknown and $k_0=[\tau_0 T]$ for some $\tau_0\in(0,1)$. The prechange mean of X_{it} is μ_{i1}, and postchange mean is μ_{i2}. The difference $\lambda_i=\mu_{i2}-\mu_{i1}$ represents the magnitude of change in ith panel, which can be either random or nonrandom, and is assumed to be independent of error process ε_{it}. In this paper, we assume that each series has common change point at k_0. Our purpose is to give consistent estimators with lower time complexity.

For a given k such that $1\leqslant k\leqslant T-1$, define

$$\overline{X}_{i1}=\frac{1}{k}\sum_{t=1}^{k}X_{it},\qquad \overline{X}_{i2}=\frac{1}{T-k}\sum_{t=k+1}^{T}X_{it}. \tag{2}$$

So, $\overline{X}_{i1}$ and $\overline{X}_{i2}$ are estimators for μ_{i1} and μ_{i2}, respectively. The classical least square estimator for k_0 in Peštová and Pešta [11] is defined as

$$\widehat{k}_{LS}=\underset{1\le k\le T-1}{\arg\min}\, SSR(k), \tag{3}$$

where

$$SSR(k)=\sum_{i=1}^{N}S_{iT}(k),\qquad S_{iT}(k)=\sum_{t=1}^{k}\left(X_{it}-\overline{X}_{i1}\right)^2+\sum_{t=k+1}^{T}\left(X_{it}-\overline{X}_{i2}\right)^2. \tag{4}$$

This estimator is straightforward to compute, and the time complexity is $O(NT^2)$. When T and N are huge, it is not to easy to locate the change point. In this paper, a two-stage estimation method is proposed to reduce the time complexity.

We adopt some assumptions in Bai [7].

Assumption 1. $\varepsilon_{it}=\sum_{j=0}^{\infty}a_{ij}e_{i,t-j}$, $e_{i,t}\sim(0,\sigma_{ie}^2)$ are i.i.d. over t; $\sum_j j|a_{ij}|\le M$ for all i. In addition, ε_{it} are independent over t. Let $\sigma_i^2=E(\varepsilon_{it})^2=\sigma_{ie}^2(\sum_j a_{ij}^2)$.

Assumption 2. $\lim_{N\longrightarrow\infty}\sum_{i=1}^{1}(\mu_{i2}-\mu_{i1})^2=\infty$.

Assumption 3. T is larger than N such that $((\log(\log(T))N)/T)\longrightarrow 0$ as T and N go to infinity.

Assumption 4. $\mu_{i2}-\mu_{i1}=N^{-1/2}\Delta_i$, with $\lim_{N\longrightarrow\infty}\sum_{i=1}^{N}(\mu_{i2}-\mu_{i1})^2=\lim_{N\longrightarrow\infty}(1/N)\sum_{i=1}^{N}\triangle_i^2=\lambda$, $\lim_{N\longrightarrow\infty}\sum_{i=1}^{N}[(\mu_{i2}-\mu_{i1})^2\ \sigma_i^2]=\lim_{N\longrightarrow\infty}(1/N)\sum_{i=1}^{N}\triangle_i^2\sigma_i^2=\phi$.

3. Two-Stage Estimator

3.1. The Initial Estimator. For any given $1\le i_0\le N$, a univariate change point series with T observations X_{i_0t} is selected from N panels. The initial estimator for k_0 is defined as

$$\widehat{l}=\underset{1\le l\le T-1}{\arg\min}\left\{\sum_{t=1}^{l}\left(X_{i_0t}-\overline{X}_{i_01}\right)^2+\sum_{t=l+1}^{T}\left(X_{i_0t}-\overline{X}_{i_02}\right)^2\right\}=\underset{1\le l\le T-1}{\arg\max}\left\{\frac{l(T-l)}{T^2}\right\}^{1/2}\left|\overline{X}_{i_02}-\overline{X}_{i_01}\right|. \tag{5}$$

Actually, the initial change point estimator $\widehat{l}$ is an ordinary least square estimator for change point in mean of univariate series. Let $\widehat{\gamma}=(\widehat{l}/T)$ and $\lambda_{i_0}=\mu_{i_02}-\mu_{i_01}$. Then, according to Proposition 3 and Theorem 1 in Bai [12], the following conclusions can be drawn.

$$\widehat{\gamma}-\tau_0=O_p\left(\frac{1}{T\lambda_{i_0}^2}\right), \tag{6}$$

$$T\lambda_{i_0}^2(\widehat{\gamma}-\tau_0)\overset{d}{\longrightarrow}\left(\sum_{j=0}^{\infty}a_{i_0j}\right)^2\sigma_{i_0}^2\arg\max_{v}\left\{W(v)-\frac{|v|}{2},\right\}, \tag{7}$$

where $W(v)$ is a two-sided Brownian motion on R.

Denote $\widehat{\lambda}_{i_0}^2=\overline{X}_{i_02}-\overline{X}_{i_01}$; $\widehat{A}_{i_0}$ is a consistent estimator for $D(\varepsilon_{i_0t})=(\sum_{j=0}^{\infty}a_{i_0j})^2\sigma_{i_0}^2$, and c_α is the α quantile of $\arg\max_v\{W(v)-|v|/2\}$. Using (7), the $(1-\alpha)\%$ confidence interval for k_0 is conducted as

$$\left[\widehat{l}-\frac{\widehat{A}_{i_0}}{\widehat{\lambda}_{i_0}^2}\cdot c_\alpha,\widehat{l}+\frac{\widehat{A}_{i_0}}{\widehat{\lambda}_{i_0}^2}\cdot c_\alpha\right]. \tag{8}$$

Given $B_T\longrightarrow\infty\,(T\longrightarrow\infty)$ and $(B_T/T)\longrightarrow 0$, we enlarge confidence interval (8) to

$$\left[\widehat{l}-B_T\cdot\frac{\widehat{A}_{i_0}}{\widehat{\lambda}_{i_0}^2}\cdot c_\alpha,\widehat{l}+B_T\frac{\widehat{A}_{i_0}}{\widehat{\lambda}_{i_0}^2}\cdot c_\alpha\right]. \tag{9}$$

Define $h_1=\widehat{l}-B_T\cdot(\widehat{A}_{i_0}/\widehat{\lambda}_{i_0}^2)\cdot c_\alpha$, $h_2=\widehat{l}+B_T\cdot(\widehat{A}_{i_0}/\widehat{\lambda}_{i_0}^2)\cdot c_\alpha$, and $T_0=h_2-h_1+1$. So,

$$P\{k_0 \in [h_1, h_2]\} \longrightarrow 1, \quad \text{as } T \longrightarrow \infty. \tag{10}$$

The time complexity for initial estimator is $O(T^2)$.

3.2. The Finial Estimator. Using all samples falling into the random interval $[h_1, h_2]$, the finial estimator is defined as

$$\widehat{k}_{TS} = \arg\min_{h_1 \le k < h_2} SSR_1(k), \tag{11}$$

where

$$SSR_1(k) = \sum_{i=1}^{N} \left(\sum_{t=h_1}^{k} (X_{it} - \overline{X}_{i1})^2 + \sum_{t=k+1}^{h_2} (X_{it} - \overline{X}_{i2})^2 \right). \tag{12}$$

Denote

$$U_{NT}(k) = \frac{1}{NT} SSR_1(k) = \frac{1}{NT} \sum_{i=1}^{N} S_{iT}(k), \tag{13}$$

and then

$$\widehat{k}_{TS} = \arg\min_{k \in [h_1, h_2]} SSR_1(k) = \arg\min_{k \in [h_1, h_2]} U_{NT}(k). \tag{14}$$

It is easy to see that time complexity for finial estimator is $O(NB_T^2)$). So, total time complexity for the two-stage estimation method is $O(T^2 + NB_T^2)$), which is smaller than $O(NT^2)$ because $(B_T/T) \longrightarrow 0$. It implies that the two-stage estimation method can give the change point position faster. Furthermore, Theorems 1 and 2 in Section 4 ensure the accuracy of our estimator.

In order to prove the properties of the two-stage estimator, we need the following lemmas described in Bai [7].

Lemma (A.1). *Assume that model (1) and Assumption 1 hold, and we have*

$$\sup_{1 \le k \le T} \left| U_{NT}(k) - EU_{NT}(k) \right| = O_p\left(\frac{1}{\sqrt{NT}} \right). \tag{15}$$

Lemma (A.2). *Assume that model (1) and Assumption 1 hold. For all $k \in [1, T]$, the expected value of $U_{NT}(k)$ satisfies*

$$EU_{NT}(k) - EU_{NT}(k_0) \ge \frac{\lambda C |k - k_0|}{(NT)}, \tag{16}$$

where $\lambda = \sum_{i=1}^{N} (\mu_{i2} - \mu_{i1})^2$ for some $C > 0$.

4. Theorem and Proof

The following properties can be obtained.

Theorem 1. *Under Assumptions 1–3, we have*

$$\lim_{N,T \longrightarrow \infty} P\left(\widehat{k}_{TS} = k_0\right) = 1. \tag{17}$$

Proof. Due to symmetry, it is sufficient to consider $k \le k_0$. According to Lemmas (A.1) and (A.2), for any $k \in [h_1, h_2]$, since

$$\begin{aligned} U_{NT}(k) - U_{NT}(k_0) &= U_{NT}(k) - EU_{NT}(k) - \left[U_{NT}(k_0) - EU_{NT}(k_0)\right] + EU_{NT}(k) - EU_{NT}(k_0) \\ \left|U_{NT}(k)\right| - \left|U_{NT}(k_0)\right| &\ge -2 \sup_{h_1 \le k \le h_2} \left|U_{NT}(k) - EU_{NT}(k)\right| + \left|EU_{NT}(k)\right| - \left|EU_{NT}(k_0)\right| \\ &\ge -2 \sup_{h_1 \le k \le h_2} \left|U_{NT}(k) - EU_{NT}(k)\right| + \frac{\lambda C |k - k_0|}{(NT)}, \end{aligned} \tag{18}$$

and $P\left\{U_{NT}(\widehat{k}_{TS}) - U_{NT}(k_0) \le 0\right\} \longrightarrow 1$, then we have that

$$\begin{aligned} \left|\widehat{k}_{TS} - k_0\right| &\le 2\lambda^{-1} C^{-1} NT \sup_{h_1 \le k \le h_2} \left|U_{NT}(k) - EU_{NT}(k)\right| \\ &= 2\lambda^{-1} C^{-1} (NT)^{1/2} O_p(1). \end{aligned} \tag{19}$$

It can be concluded from Assumptions 1 and 2 and (19) that

$$\frac{\left|\widehat{k}_{TS} - k_0\right|}{T} = \frac{2}{\lambda C} \sqrt{\frac{N}{T}} O_p(1) = o_p(1). \tag{20}$$

Thus, $|\widehat{k}_{TS} - k_0| = o_p(T)$. So, for any $\varepsilon > 0$, $|\widehat{k}_{TS} - k_0| \le \varepsilon T$ for large T, with probability tending to 1. Because $k_0 = [T\tau_0]$ and $\tau \in (0,1)$, there exists $\delta > 0$ such that $P(\widehat{k}_{TS} \in [\delta T, (1 - \delta)T]) \longrightarrow 1$ as $T \longrightarrow \infty$. That is, $P(\widehat{k}_{TS} \in D) \longrightarrow 1$, where $D = \{k\colon \delta T \le k \le (1-\delta)T, h_1 \le k \le h_2\}$.

Define the set $D(k_0) = D/\{k_0\}$ so that $D(k_0)$ excludes k_0 from D. Then,

$$\begin{aligned} P\left(\widehat{k}_{TS} \ne k_0\right) &\le P\left(\widehat{k}_{TS} \notin D\right) + P\left(\widehat{k}_{TS} \in D, \widehat{k}_{TS} \ne k_0\right) \\ &= P\left(\widehat{k}_{TS} \notin D\right) + P\left(\widehat{k}_{TS} \in D(k_0)\right). \end{aligned} \tag{21}$$

By the definition of $\widehat{k}_{TS}$, $U_{NT}(\widehat{k}_{TS}) \le U_{NT}(k_0)$. So, a necessary condition for $\widehat{k}_{TS} \in D(k_0)$ is $\min_{k \in D(k_0)} U_{NT}(k) U_{NT}(k_0) \le 0$. Similar to Lemma A.3 of Bai [7], it can be proved that $P\left\{\min_{k \in D(k_0)} U_{NT}(k) - U_{NT}(k_0) \le 0\right\} \longrightarrow 0$, which implies that $P(\widehat{k}_{TS} \in D(k_0)) \longrightarrow 0$. Thus,

$$P\left(\widehat{k}_{TS} \ne k_0\right) \le P\left(\widehat{k}_{TS} \notin D\right) + P\left(\widehat{k}_{TS} \in D(k_0)\right) \longrightarrow 0. \tag{22}$$

This completes the proof. □

Theorem 2. *Under Assumptions 1, 3, and 4, as $T, N \longrightarrow \infty$,*

$$\widehat{k}_{TS} - k_0 \xrightarrow{d} \arg\min_{\ell}[|\ell|\lambda + 2\sqrt{\phi}W(\ell)], \tag{23}$$

where $W(0) = 0$ and $W(\ell) = \sum_{s=-\ell+1}^{0} Z_s, \ell = -1, -2, \ldots$, $W(\ell) = \sum_{s=1}^{\ell} Z_s$, $\ell = 1, 2, \ldots$, and $Z_s, s = \ldots, -1, 0, 1, \ldots$ are i.i.d. standard normal random variables.

Proof. Again, by symmetry, it is sufficient to consider $k \leq k_0$. For $k \leq k_0$ and $k \in [h_1, h_2]$,

$$\begin{aligned}\overline{X}_{i1} &= \mu_{i1} + \frac{1}{k}\sum_{t=h_1}^{k}\varepsilon_{it},\\ \overline{X}_{i2} &= \mu_{i1} + \frac{T_0 - k_0}{T_0 - k}(\mu_{i2} - \mu_{i1}) + \frac{1}{T_0 - k}\sum_{t=k+1}^{h_2}\varepsilon_{it}\\ &= \frac{k_0 - k}{T_0 - k}(\mu_{i1} - \mu_{i2}) + \mu_{i2} + \frac{1}{T_0 - k}\sum_{t=k+1}^{h_2}\varepsilon_{it}.\end{aligned} \tag{24}$$

Introduce

$$\begin{aligned}\overline{\varepsilon}_{i1} &= \frac{1}{k}\sum_{t=h_1}^{k}\varepsilon_{it},\\ \overline{\varepsilon}_{i2} &= \frac{1}{T_0 - k}\sum_{t=k+1}^{h_2}\varepsilon_{it},\end{aligned} \tag{25}$$

and let

$$\begin{aligned}a_{ik} &= \frac{T_0 - k_0}{T_0 - k}(\mu_{i2} - \mu_{i1}),\\ b_{ik} &= \frac{k_0 - k}{T_0 - k}(\mu_{i1} - \mu_{i2}).\end{aligned} \tag{26}$$

It follows that

$$\begin{aligned}\overline{X}_{i1} &= \mu_{i1} + \overline{\varepsilon}_{i1},\\ \overline{X}_{i2} &= \mu_{i1} + a_{ik} + \overline{\varepsilon}_{i2} = \mu_{i2} + b_{ik} + \overline{\varepsilon}_{i2}.\end{aligned} \tag{27}$$

By the definition of $S_{iT}(k)$, we get

$$\begin{aligned}S_{iT}(k) &= \sum_{t=h_1}^{k}(\varepsilon_{it} - \overline{\varepsilon}_{i1})^2 + \sum_{t=k+1}^{k_0}(\varepsilon_{it} - a_{ik} - \overline{\varepsilon}_{i2})^2\\ &\quad + \sum_{t=k_0+1}^{h_2}(\varepsilon_{it} - b_{ik} - \overline{\varepsilon}_{i2})^2\\ &= \sum_{t=h_1}^{k}(\varepsilon_{it} - \overline{\varepsilon}_{i1})^2 + \sum_{t=k+1}^{h_2}(\varepsilon_{it} - \overline{\varepsilon}_{i2})^2\\ &\quad + (k_0 - k)a_{ik}^2 + (T_0 - k)b_{ik}^2 - 2a_{ik}\sum_{t=k+1}^{k_0}(\varepsilon_{it} - \overline{\varepsilon}_{i2})\\ &\quad - 2b_{ik}\sum_{t=k_0+1}^{h_2}(\varepsilon_{it} - \overline{\varepsilon}_{i2}).\end{aligned} \tag{28}$$

Notice that

$$\sum_{t=h_1}^{k}(\varepsilon_{it} - \overline{\varepsilon}_{i1})^2 + \sum_{t=k+1}^{h_2}(\varepsilon_{it} - \overline{\varepsilon}_{i2})^2 = \sum_{t=h_1}^{h_2}\varepsilon_{it}^2 - k\overline{\varepsilon}_{i1}^2 - (T_0 - k)\overline{\varepsilon}_{i2}^2, \tag{29}$$

and thus

$$\begin{aligned}SSR_1(k) &= (k_0 - k)\sum_{i=1}^{N}a_{ik}^2 + (T - k_0)\sum_{i=1}^{N}b_{ik}^2 + \sum_{i=1}^{N}\sum_{t=h_1}^{h_2}\varepsilon_{it}^2\\ &\quad - \sum_{i=1}^{N}k\overline{\varepsilon}_{i1}^2 - \sum_{i=1}^{N}(T_0 - k)\overline{\varepsilon}_{i2}^2\\ &\quad - 2\sum_{i=1}^{N}a_{ik}\sum_{t=k+1}^{k_0}(\varepsilon_{it} - \overline{\varepsilon}_{i2}) - 2\sum_{i=1}^{N}b_{ik}\sum_{t=k_0+1}^{h_2}(\varepsilon_{it} - \overline{\varepsilon}_{i2}).\end{aligned} \tag{30}$$

So,

$$\begin{aligned}SSR_1(k) - SSR_1(k_0) &= (k_0 - k)\sum_{i=1}^{N}a_{ik}^2 + (T - k_0)\sum_{i=1}^{N}b_{ik}^2 - \sum_{i=1}^{N}\left[\frac{1}{k}\left(\sum_{t=h_1}^{k}\varepsilon_{it}\right)^2 - \frac{1}{k_0}\left(\sum_{t=h_1}^{k_0}\varepsilon_{it}\right)^2\right]\\ &\quad - \sum_{i=1}^{N}\left[\frac{1}{T_0 - k}\left(\sum_{t=k+1}^{h_2}\varepsilon_{it}\right)^2 - \frac{1}{T_0 - k_0}\left(\sum_{t=k_0+1}^{h_2}\varepsilon_{it}\right)^2\right]\\ &\quad - 2\sum_{i=1}^{N}a_{ik}\sum_{t=k+1}^{k_0}\varepsilon_{it} - 2\sum_{i=1}^{N}b_{ik}\sum_{t=k_0+1}^{h_2}\varepsilon_{it} + 2\sum_{i=1}^{N}[(k_0 - k)a_{ik} + (T_0 - k_0)b_{ik}]\overline{\varepsilon}_{i2}.\end{aligned} \tag{31}$$

Table 1: Simulation comparison of two-stage estimation and least square estimation.

		$N = 80, T = 1000$		$N = 100, T = 1200$		$N = 120, T = 1400$	
		Mean (Sd)	$Tc\,(s)$	Mean (Sd)	$Tc\,(s)$	Mean (Sd)	$Tc\,(s)$
	$\lg(T)$	501.218 (26.411)	0.926	601.084 (21.981)	1.180	700.164 (20.535)	1.468
TS	$\ln(T)$	499.998 (0.184)	2.452	600.012 (0.126)	3.194	700.006 (0.073)	3.838
	$\log_2(T)$	499.998 (0.118)	3.834	600.006 (0.134)	5.074	700.000 (0.063)	5.946
LS		499.998 (0.195)	7.640	599.988 (0.141)	11.378	699.998 (0.077)	24.856

According to Theorem 4.2 of Bai [7], the first and the fifth terms are $(k_0 - k)O_p(1)$ and all others are $(k_0 - k)O_p(1)$.

Because $((T_0 - k_0)/(T_0 - k)) \xrightarrow{p} 1$, we have

$$(k_0 - k)\sum_{i=1}^{N} a_{ik}^2 = (k_0 - k)\sum_{i=1}^{N}\left\{\frac{T_0 - k_0}{T_0 - k}(\mu_{i2} - \mu_{i1})\right\}^2 \xrightarrow{p} (k_0 - k)\lambda. \tag{32}$$

Similarly,

$$\begin{aligned} -2\sum_{i=1}^{N} a_{ik}\sum_{t=k+1}^{k_0}\varepsilon_{it} &= -2\frac{T_0 - k_0}{T_0 - k}\sum_{t=k+1}^{k_0}\left\{\sum_{i=1}^{N}(\mu_{i2} - \mu_{i1})\varepsilon_{it}\right\} \\ &= -2\frac{T_0 - k_0}{T_0 - k}\sum_{t=k+1}^{k_0} N^{-1/2}\sum_{i=1}^{N}\Delta_i\varepsilon_{it} = -2\frac{T_0 - k_0}{T_0 - k}\sqrt{\phi_N}\sum_{t=k+1}^{k_0}\left(N^{-1/2}\sum_{i=1}^{N}\omega_i\varepsilon_{it}\right), \end{aligned} \tag{33}$$

where

$$\omega_i = \frac{\Delta_i}{\sqrt{\phi_N}} = \frac{\Delta_i}{\left((1/N)\sum_{j=1}^{N}\Delta_j^2\sigma^2\right)^{1/2}}. \tag{34}$$

Under Assumption 4, $\phi_N = (1/N)\sum_{j=1}^{N}\Delta_j^2\sigma_j^2 \longrightarrow \phi$ and $N^{-1/2}\sum_{i=1}^{N}\omega_i\varepsilon_{it} \longrightarrow Z_t$, where $Z_t \sim N(0,1)$. Thus, the limit of the fifth term of (31) is $2\phi^{1/2}\sum_{t=k+1}^{k_0} Z_t$.

In summary, for $k \le k_0$,

$$SSR_1(k) - SSR_1(k_0) \xrightarrow{d} (k_0 - k)\lambda + 2\phi^{1/2}\sum_{t=k+1}^{k_0} Z_t. \tag{35}$$

Similarly, for $k > k_0$, we can prove that

$$SSR_1(k) - SSR_1(k_0) \xrightarrow{d} (k - k_0)\lambda + 2\phi^{1/2}\sum_{t=k_0+1}^{k} Z_t. \tag{36}$$

Let $\ell = k_0 - k$, $\widehat{k}_{TS} - k_0 \xrightarrow{d} \arg\min_{\ell}[|\ell|\lambda + 2\sqrt{\phi}W(\ell)]$ by functional central limit theorem. □

5. Monte Carlo Comparison

We compare two estimators (3) and (11) by Monte Carlo simulation on the same computer. The series is generated according to model (1), where $\varepsilon_{it} \sim N(0,1)$, $\mu_{i1} \sim U(-10,10)$ and $\mu_{i1} - \mu_{i2} \sim N(0.5, 0.01)$. Experiments are carried out for $N = 80, 100, 120$, $T = 1000, 1200, 1400$, and $k_0 = 0.5T$. We choose $B_T = \lg(T)$, $\ln(T)$, $\log_2(T)$. Table 1 reports our simulation results based on 500 replications, where Mean, $S\,d$, and Tc stand for the average estimator, standard deviation, and the operated time for the computer, respectively, while time is in seconds.

It can be seen from Table 1 that with the increasing number of sequences of N and the sample size of T, the running times are getting longer and longer for both the two-stage method and least square method. When N and T are fixed, with the increase of B_T, the two-stage estimator is closer to the true change point and the running time is increased slightly. The running time of the two-stage estimation method is much less than that of the least square method. This shows that in the case of massive data, the method in this paper can estimate the change point position faster.

Data Availability

All data are computer simulation data.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (grant nos. 11771353 and 12171391).

References

[1] L. Joseph and D. B. Wolfson, "Estimation in multi-path change-point problems," *Communications in Statistics-Theory and Methods*, vol. 21, no. 4, pp. 897–913, 1992.

[2] L. Joseph and D. B. Wolfson, "Maximum likelihood estimation in the multi-path change-point problem," *Annals of the Institute of Statistical Mathematics*, vol. 45, no. 3, pp. 511–530, 1993.

[3] L. Joseph, A. C. Vandal, and D. B. Wolfson, "Estimation in the multipath change point problem for correlated data," *Canadian Journal of Statistics*, vol. 24, no. 1, pp. 37–53, 1996.

[4] L. Joseph, D. B. Wolfson, R. D. Berger, and R. M. Lyle, "Analysis of panel data with change-points," *Statistica Sinica*, vol. 7, pp. 687–703, 1997.

[5] S. J. Skates, D. K. Pauler, and I. J. Jacobs, "Screening based on the risk of cancer calculation from bayesian hierarchical changepoint and mixture models of longitudinal markers," *Journal of the American Statistical Association*, vol. 96, no. 454, pp. 429–439, 2001.

[6] C. H. Jackson and L. D. Sharples, "Models for longitudinal data with censored change points," *Applied Statistcs*, vol. 24, pp. 37–53, 2004.

[7] J. Bai, "Common breaks in means and variances for panel data," *Journal of Econometrics*, vol. 157, no. 1, pp. 78–92, 2010.

[8] L. Horváth and M. Hušková, "Change-point detection in panel data," *Journal of Time Series Analysis*, vol. 33, no. 4, pp. 631–648, 2012.

[9] D. W. Shin and E. Hwang, "A CUSUM test for panel mean change detection," *Journal of the Korean Surgical Society*, vol. 46, no. 1, pp. 70–77, 2017.

[10] P. Cao and Z. Xia, "A fast estimation method for mean change point in massive data sets," *Chinese Journal of Applied Probability and Statistics*, vol. 36, no. 5, pp. 493–508, 2020.

[11] B. Peštová and M. Pešta, "Testing structural changes in panel data with small fixed panel size and bootstrap," *Metrika*, vol. 78, pp. 665–689, 2015.

[12] J. Bai, "Least squares estimation of a shift in linear processes," *Journal of Time Series Analysis*, vol. 15, no. 5, pp. 453–472, 1994.