**ISSN** 1749-3889 (print), 1749-3897 (online)
**International Journal of Nonlinear Science**
Vol.35(2023) No.1,pp.24-30

# A Consistent Estimator for the Change Point in AR(1) Models: $I(0) - I(1)$ Process

Zongxian Zhao*
School of Mathematical Sciences, Jiangsu University, Zhenjiang, Jiangsu 212013, P. R. China
*(Received November 18 2022, accepted February 13 2023)*

**Abstract:** Change-point detection is an important topic in the time series analysis. For AR(1) models, the change of parameters in the model may lead different properties, e.g., stationarity. This paper studies AR(1) models with single change point, where it is a stationary process $I(0)$ before the change and a non-stationary unit root process $I(1)$ after the change. The data generated process is refined to avoid the spurious jump at the true change point. We propose a least squares estimator of the change point and show that it converges to the true value in probability at the rate strictly $T$. This result is symmetric with the convergence rate in the process from $I(1)$ to $I(0)$. The R language is used to implement numerical simulations that illustrate our theoretical findings.

**Keywords:** Change-point problem; Autoregression model; Consistent estimator; Non-stationary process

## 1 Introduction

The change-point problems has attracted extensive attention in the area of statistics, economics and finance in recent years. In reality, observed data are usually generated from more than one latent status, e.g., bullish and bearish markets in finance, prosperity and recession in economic data (Frick et al. (2014) [1]). The detection of change points is critical to identify the status where we are living in.

The autoregression (AR) model is a standard model in the time series analysis and has widely applications in finance and economics. The AR(1) model can be written as follows:

$$y_t = \beta y_{t-1} + \epsilon_t, \;\; t = 1, \cdots, T,$$

where $\epsilon_t$ are i.i.d random errors with zero means and finite variances $\sigma^2$. The structure change in the AR(1) model means that the parameter $\beta$ is different before and after some unknown time $k_0$, which is called a change point. Mankiw et al. (1987) [2] explored the structure change in U.S. short-term interest rate using AR(1) models. Other applications include Burdekin and Siklos (1999) [3] for the inflation rate series, Phillips et al. (2011) [4] and Phillips and Shi (2018) [5] for detecting the boom and burst of financial bubbles, and so on.

It is well known that the AR(1) model is stationary if and only if the absolute value of $\beta$ is less than unity. So the changes of $\beta$ may affect the stationarity of models. Chong (2001) [6] proposed a generalized AR(1) model with single change point:

$$y_t = \beta_1 y_{t-1}\delta\{t \le k_0\} + \beta_2 y_{t-1}\delta\{t > k_0\} + \epsilon_t, \;\; t = 1, \cdots, T, \tag{1}$$

where $\delta\{\cdot\}$ denotes the indicator function and $k_0$ is the true change point. They explored the consistency and limiting distributions of the least squares estimators of $\beta_1, \beta_2$ and unknown fraction break $\lambda_0 (= k_0/T)$ for the following three cases: (i) $|\beta_1| < 1$ and $|\beta_2| < 1$ (from a stationary process $I(0)$ to $I(0)$); (ii) $|\beta_1| < 1$ and $\beta_2 = 1$ (from $I(0)$ to a non-stationary unit root process $I(1)$); (iii) $\beta_1 = 1$ and $|\beta_2| < 1$ (from $I(1)$ to $I(0)$). In all of the cases, Chong (2001) [6] showed the least squares estimator $\hat{\lambda}$ was always consistent with the true value, but the rates of convergence differed from case to case. Specifically, they found the convergence rate in the $I(1) - I(0)$ process is faster than that in the $I(0) - I(1)$ process. Many researches follow the framework in Equation 1, e.g., Pang and Zhang (2015) [7], Pang et al. (2018), [8], Mohr and Selk (2020) [9]. However, Kejriwal and Perron (2012) [10] explored the reasons why the rates of convergence

***Corresponding author**. *E-mail address*: jsuzhaozx@163.com


IJNS.2023.03.15

are different between $I(0) - I(1)$ and $I(1) - I(0)$ cases. They pointed out that the data generation process in Equation 1 imposed a spurious jump at the change point for the $I(1) - I(0)$ case. To avoid this spurious jump, they proposed an improved data generation process for the $I(1) - I(0)$ case. Under the improved data generation process, the estimator of break fraction $\hat{\lambda}$ was shown to converge to the true value $\lambda_0$ in probability at the rate $T$.

In this paper, we extend the framework in Kejriwal and Perron (2012) [10] to the $I(0) - I(1)$ case. We show that the least squares estimator of break fraction $\hat{\lambda}$ also converges in probability to the true value $\lambda_0$ at the rate strictly $T$ for the $I(0)-I(1)$ case. We fix the spurious jump and asymmetry problem in Chong (2001) [6], and unify both cases of structure changes into one framework. Monte Carlo simulations with R language are also conducted to demonstrate our theoretical results in finite samples.

This paper is organized as follows. Section 2 discusses the spurious jump problem in Chong (2001) [6] and formulate our data generate process in the $I(0) - I(1)$ case. Section 3 proposes a least square estimator $\hat{\lambda}$ and shows its asymptotic properties. Section 4 presents simulation studies to support our theoretical results. Section 5 concludes the paper.

# 2 Models

## 2.1 Spurious Jumps in the I(1)-I(0) Process

Chong (2001) [6] proposed a generalized AR(1) model with single change point in Equation 1. For the $I(1) - I(0)$ case, it can be written as:

$$
\begin{aligned}
y_t &= y_{t-1} + \epsilon_t, \;\; t = 1, \cdots, k_0, \\
y_t &= \beta_2 y_{t-1} + \epsilon_t, \;\; t = k_0 + 1, \cdots, T,
\end{aligned} \tag{2}
$$

where $k_0 = [\lambda_0 T]$ is the true change point. Here $[x]$ means that the fraction $x$ is rounded down. The model also has the following assumptions:

**Assumption 1**: $y_0$ is a random variable with zero mean and finite variance;
**Assumption 2**: $\epsilon_t, t = 1, \cdots, T$, are i.i.d random errors with zero means and finite variances $\sigma^2$, where $E(\epsilon_t^4) < \infty$;
**Assumption 3**: $\lambda_0 \in (0, 1)$;
**Assumption 4**: $|\beta_2| < 1$.

However, according to the data generation process in Equation 2, given $y_0, y_1, \cdots, y_{k_0}$, we have

$$
y_t = \beta_2^{t-k_0} y_{k_0} + \sum_{j=0}^{t-k_0-1} \beta_2^j \epsilon_{t-j}, \;\; t = k_0 + 1, \cdots, T.
$$

The conditional expectation is

$$
E(y_t | y_0, \cdots, y_{k_0}) = \beta_2^{t-k_0} y_{k_0} \to 0 \;\text{ as }\; t \to \infty
$$

since $|\beta_2| < 1$. It means the $I(0)$ process always converges to zero no matter how large what the value of $y_{k_0}$ is. The smaller $\beta_2$ leads to a faster convergence. It generates a spurious jump at the change point. Figure 1 presents the problem under different $\beta_2$.

To avoid the spurious jump problem in the $I(1) - I(0)$ case, Kejriwal and Perron (2012) [10] proposed a revised data generation process as follows:

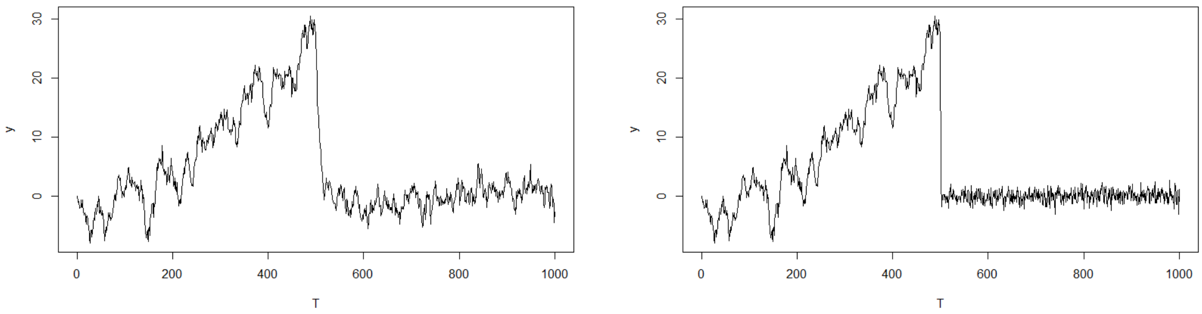
$$
\begin{aligned}
y_t &= y_{t-1} + \epsilon_t, \;\; t = 1, \cdots, k_0, \\
y_t &= y_{k_0} + \beta(y_{t-1} - y_{k_0}) + \epsilon_t, \;\; t = k_0 + 1, \cdots, T,
\end{aligned} \tag{3}
$$

where $k_0 = [\lambda_0 T]$ is the true change point. This model also follows the above Assumption 2-4, and it substitutes Assumption 1 by Assumption 5 as follows:

**Assumption 5**: $y_0$ is a constant or an $O_p(1)$ random variable, i.e., it is bounded in probability (tight).

Based on the data generation process in Equation 3, we have

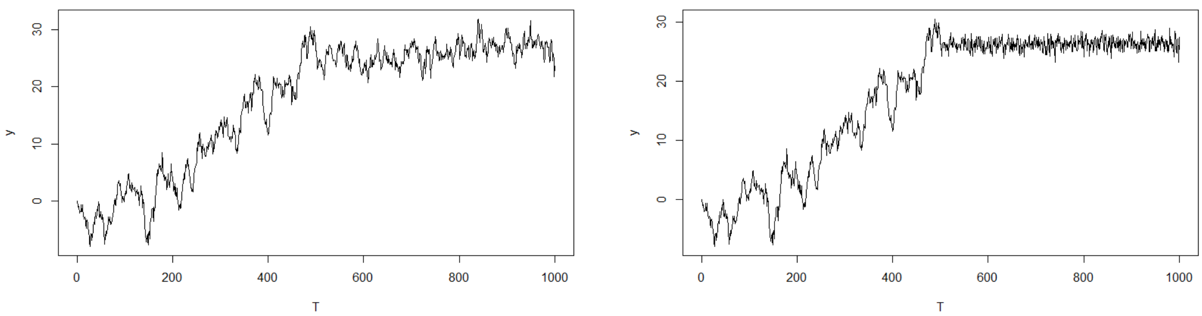
$$
y_t = y_{k_0} + \sum_{j=0}^{t-k_0-1} \beta_2^j \epsilon_{t-j}, \;\; t = k_0 + 1, \cdots, T.
$$




(a) $\beta_2 = 0.9, T = 1000, k_0 = 500$

(b) $\beta_2 = 0.1, T = 1000, k_0 = 500$

Figure 1: (a) and (b) demonstrate the spurious jump problem at the change point for the $I(1) - I(0)$ case when data are generated from Equation 2. The smaller $\beta_2$ in (b) leads to a faster convergence than that in (a).




(a) $\beta = 0.9, T = 1000, k_0 = 500$

(b) $\beta = 0.1, T = 1000, k_0 = 500$

Figure 2: The data generation process in Equation 3. Compared with the data generation process in Equation 2 (Figure 1), the spurious jump problem at the change point is fixed.

Then the conditional expectation $E(y_t|y_0, \cdots, y_{k_0}) = y_{k_0}$ is a constant without the spurious jump. Figure 2 presents the data generation process in Equation 3.

Under the data generation process in Equation 3 and model assumptions, Kejriwal and Perron (2012) [10] showed the estimator of break fraction $\hat{\lambda}_0$ would converge to the true value $\lambda_0$ in probability at the rate $T$ for the $I(1) - I(0)$ case.

### 2.2 Models for the I(0)-I(1) Process

For the $I(0) - I(1)$ case, the data generation process in Chong (2001) [6] can be written as

$$\begin{aligned} y_t &= \beta_1 y_{t-1} + \epsilon_t, \;\; t = k_0 + 1, \cdots, T, \\ y_t &= y_{t-1} + \epsilon_t, \;\; t = 1, \cdots, k_0, \end{aligned} \tag{4}$$

where $k_0 = [\lambda_0 T]$ is the true change point and $|\beta_1| < 1$. Note that

$$y_t = \beta_1^t y_0 + \sum_{j=0}^{t-1} \beta_1^j \epsilon_{t-j}, \;\; t = 1, \cdots, k_0.$$

Given $y_0$, the conditional expectation is $E(y_t|y_0) = \beta_1^t y_0$. If $y_0 \neq 0$, then $E(y_t|y_0)$ decays to zero as $t$ gets larger. This also causes a spurious jump at the beginning point $y_0$.

To avoid this drawback, we propose a revised data generation process inspired by that in Kejriwal and Perron (2012) [10]. The data generation process for the $I(0) - I(1)$ case can be expressed as

$$\begin{aligned} y_t &= y_0 + \alpha(y_{t-1} - y_0) + \epsilon_t, \;\; t = 1, \cdots, k_0, \\ y_t &= y_{t-1} + \epsilon_t, \;\; t = k_0 + 1, \cdots, T, \end{aligned} \tag{5}$$

where $k_0 = [\lambda_0 T]$ is the true change point. It follows the Assumption 2-5 with the only difference that it requires $|\alpha| < 1$ rather than $|\beta_2| < 1$ in Assumption 4. Then, we have

$$y_t = y_0 + \sum_{j=0}^{t-1} \alpha^j \epsilon_{t-j}, \;\; E(y_t|y_0) = y_0, \tag{6}$$

where $t = 1, \cdots, k_0$. There is no spurious jump in the data generation process.

## 3 Estimator of $\lambda_0$ and Asymptotic Properties

### 3.1 Notations

We define the notations used in the following analyses for convenience.

$$\begin{aligned} & k = [\lambda T], \;\; k_0 = [\lambda_0 T], \\ & \bar{y}_1 = \frac{1}{k}\sum_{t=1}^{k} y_t, \;\; \bar{y}_{1,-1} = \frac{1}{k}\sum_{t=1}^{k} y_{t-1}, \;\; \tilde{y}_{1,t-1} = y_{t-1} - \bar{y}_{1,-1}, \\ & \bar{y}_2 = \frac{1}{T-k}\sum_{t=k+1}^{T} y_t, \;\; \bar{y}_{2,-1} = \frac{1}{T-k}\sum_{t=k+1}^{T} y_{t-1}, \;\; \tilde{y}_{2,t-1} = y_{t-1} - \bar{y}_{2,-1}, \\ & z_t = \sum_{j=0}^{t-1} \alpha^j \epsilon_{t-j}, \;\; t = 1, \cdots, k_0, \;\; w_t = \sum_{j=k_0+1}^{t} \epsilon_j, \;\; t = k_0 + 1, \cdots, T. \end{aligned}$$

Thus, data generated from Equation 5 can be written as $y_t = y_0 + z_t$ for $t = 1, \cdots, k_0$ and $y_t = y_{k_0} + w_t$ for $t = k_0 + 1, \cdots, T$.

Finally, let the symbol $\xrightarrow{p}$ denotes the convergence in probability and $\xrightarrow{d}$ denotes the convergence in distribution.

### 3.2 Least Squares Estimator

The structure change in the AR(1) model means the change of coefficients in the autoregression equation. Intuitively, the ordinary least squares (OLS) method can be applied in the estimation of the change point. In this section, we propose an OLS estimator of break fraction $\lambda$ for the data generation process in Equation 5.

Let $\lambda$ be the unique break fraction, then the OLS estimator of $\alpha$ can be expressed as

$$\hat{\alpha}(\lambda) = \frac{\sum_{t=1}^{[\lambda T]}(y_t - \bar{y}_1)(y_{t-1} - \bar{y}_{1,-1})}{\sum_{t=1}^{[\lambda T]}(y_{t-1} - \bar{y}_{1,-1})^2}. \tag{7}$$

Given the estimator $\hat{\alpha}(\lambda)$, the sum of squared residuals (SSR) for the whole observations is

$$\text{SSR}(\lambda) = \sum_{t=1}^{[\lambda T]}(y_t - \bar{y}_1 - \hat{\alpha}(\lambda)(y_{t-1} - \bar{y}_{1,-1}))^2 + \sum_{t=[\lambda T]+1}^{T}(y_t - \bar{y}_2 - (y_{t-1} - \bar{y}_{2,-1}))^2.$$

The OLS estimator of $\lambda$ is defined as

$$\hat{\lambda} = \arg\min_{\lambda \in (0,1)} \text{SSR}(\lambda). \tag{8}$$

Then we show that the estimator $\hat{\lambda}$ converges to the true value $\lambda_0$ in probability at the rate strictly $T$.

### 3.3 Asymptotical Properties

First, we consider the case that $0 < \lambda < \lambda_0$.

**Lemma 1** *For $0 < \lambda < \lambda_0$, we have the following properties,*
*(i) $z_t = O_p(1)$ for $t = 1, \cdots, [\lambda_0 T]$;*
*(ii) $\sum_{t=1}^{[\lambda T]} z_t = O_p(T^{1/2})$, $\sum_{t=[\lambda T]+1}^{[\lambda_0 T]} z_t = O_p(T^{1/2})$;*
*(iii) $\sum_{t=1}^{[\lambda T]} z_t^2 = O_p(T)$, $\sum_{t=[\lambda T]+1}^{[\lambda_0 T]} z_t^2 = O_p(T)$;*
*(iv) $\sum_{t=1}^{[\lambda T]} \epsilon_t z_{t-1} = O_p(T^{1/2})$, $\sum_{t=[\lambda T]+1}^{[\lambda_0 T]} \epsilon_t z_{t-1} = O_p(T^{1/2})$.*

**Lemma 2** *For $0 < \lambda < \lambda_0$, the least squares estimator $\hat{\alpha}(\lambda)$ in Equation 7 satisfies that*

$$\hat{\alpha}(\lambda) - \alpha = o_p(1).$$

**Theorem 3** *Under the data generation process in Equation 5 and its assumptions, we have*

$$\frac{1}{T}\text{SSR}(\lambda) \xrightarrow{p} \sigma^2 + \frac{(1-\alpha)(\lambda_0 - \lambda)\sigma^2}{1+\alpha},$$

*for $0 < \lambda < \lambda_0$.*

Theorem 3 demonstrates that the $T^{-1}\text{SSR}(\lambda)$ converges to a downward sloping linear function in probability for $0 < \lambda < \lambda_0$. The $T^{-1}\text{SSR}(\lambda)$ obtains the minimum when $\lambda = \lambda_0$.

Then, we consider the case that $\lambda_0 < \lambda < 1$.

**Lemma 4** *For $\lambda_0 < \lambda < 1$, we have the following properties,*
*(i) $w_{[\lambda T]} = O_p((k-k_0)^{1/2})$, $w_T = O_p(T^{1/2})$;*
*(ii) $\sum_{t=[\lambda_0 T]+1}^{[\lambda T]} w_t = O_p((k-k_0)^{3/2})$, $\sum_{t=[\lambda T]+1}^{T} w_t = O_p(T^{3/2})$;*
*(iii) $\sum_{t=[\lambda_0 T]+1}^{[\lambda T]} w_t^2 = O_p((k-k_0)^2)$, $\sum_{t=[\lambda T]+1}^{T} w_t^2 = O_p(T^2)$;*
*(iv) $\sum_{t=[\lambda_0 T]+1}^{[\lambda T]} \epsilon_t w_{t-1} = O_p(k-k_0)$, $\sum_{t=[\lambda T]+1}^{T} \epsilon_t w_{t-1} = O_p(T)$.*

When $\lambda \in (\lambda_0, 1)$ uniformly, it leads to that $k - k_0 = [\lambda T] - [\lambda_0 T] = O(T)$. Then, we have

**Lemma 5** *For $\lambda_0 < \lambda < 1$, the least squares estimator $\hat{\alpha}(\lambda)$ in Equation 7 satisfies that*

$$\hat{\alpha}(\lambda) - 1 = O_p(1/T).$$

**Theorem 6** *Under the data generation process in Equation 5 and its assumptions, we have*

$$\frac{1}{T}SSR(\lambda) \xrightarrow{p} \frac{(1-\alpha)\lambda_0}{1+\alpha}\sigma^2 + \sigma^2,$$

*for $\lambda_0 < \lambda < 1$.*

Theorem 6 demonstrates that the $T^{-1}$SSR($\lambda$) converges to a constant greater than $\sigma^2$ in probability for $\lambda_0 < \lambda < 1$. According to above-mentioned Theorem , we can draw the following result immediately:

**Theorem 7** *The least squared estimator $\hat{\lambda}$ defined in Equation 8 converges to the true value $\lambda_0$ in probability, i.e.,*

$$\hat{\lambda} \xrightarrow{p} \lambda_0.$$

Furthermore, we explore the rate of convergence of $\hat{\lambda}$ and have Theorem 7 as follows.

**Theorem 8** *The rate of convergence of $\hat{\lambda}$ to $\lambda_0$ is strictly $T$, i.e.,*

$$T(\hat{\lambda} - \lambda_0) = O_p(1).$$

Theorem 8 shows that the least squared estimator $\hat{\lambda}$ converges to the true value at rate $T$, which is symmetric with that in [10]. It fixes the asymmetric problem of convergence rates in [6].

# 4 Simulations

Section 3 shows consistency of the least squared estimator theoretically. In this section, we presents simulation studies to demonstrate our results in finite samples.

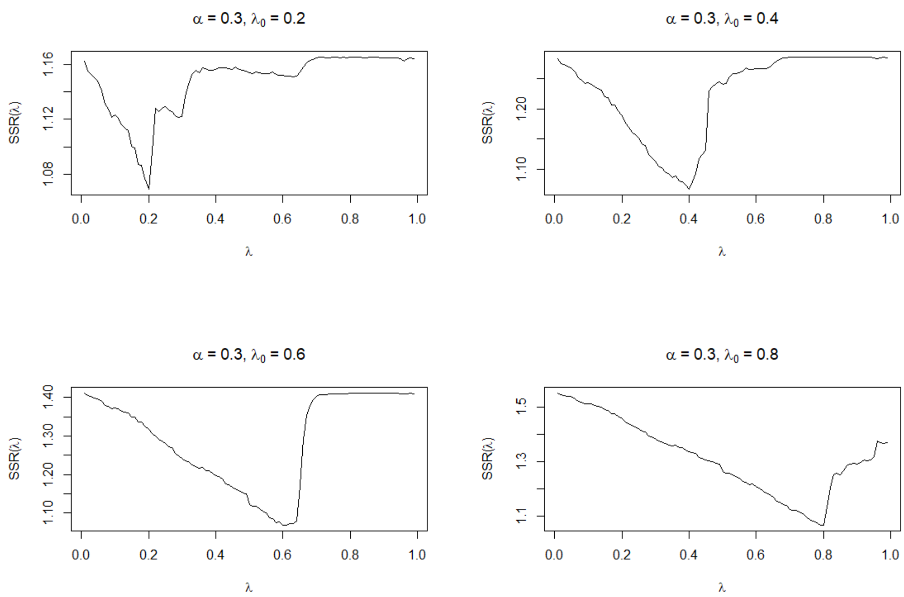



Figure 3: The plots of SSR($\lambda$) versus $\lambda$. The minimum of SSR($\lambda$) is quite close to the true value of $\lambda_0$.

Figure 3 presents the change of SSR($\lambda$) with $\alpha = 0.3$ in Table 1. The results for $\alpha = 0.5$ and 0.7 are similar and thus are omitted here. It is clear to demonstrate that the estimator $\hat{\lambda} = \arg\min_\lambda$ SSR($\lambda$) is quite close to the true value of $\lambda_0$.

# 5 Conclusions

This paper extends the model in Kejriwal and Perron (2012) [10] to the I(0)-I(1) case. The data generated process is refined to avoid the spurious jump at the true change point. We propose a least squares estimator of the change point and show that it converges to the true value in probability at the rate strictly $T$. This result is symmetric with that in [10] for the I(1)-I(0) case. Simulations with R language are conducted to demonstrate our theoretical results in finite samples. They show that our estimator is quite close to the true value of $\lambda_0$. In future works, the time series with multiple change points may be explored based on similar ideas.

# References

[1] K. Frick, A. Munk and H. Sieling. Multiscale change point inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2014;76(3):495-580.

[2] N. G. Mankiw, J. A. Miron and D. N. Weil. The adjustment of expectations to a change in regime: A study of the founding of the federal reserve. *The American Economic Review*, 1987; 77(3):358.

[3] R. C. Burdekin and P. L. Siklos. Exchange rate regimes and shifts in inflation persistence: does nothing else matter? *Journal of Money, Credit, and Banking*, 1999;235-247.

[4] P. C. Phillips, Y. Wu and J. Yu. Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review*, 2011;52(1):201-226.

[5] P. C. Phillips and S. P. Shi. Financial bubble implosion and reverse regression. *Econometric Theory*, 2018;34(4):705-753.

[6] T. T. L. Chong. Structural change in AR(1) models. *Econometric Theory*, 2001;17(1):87-155.

[7] T. Pang and D. Zhang. Asymptotic inferences for an AR(1) model with a change point and possibly infinite variance. *Communications in Statistics-Theory and Methods*, 2015; 44(22):4848-4865.

[8] T. Pang, T. T. L. Chong and D. Zhang, et al. Structural change in nonstationary AR(1) models. *Econometric Theory*, 2018;34(5):985-1017.

[9] M. Mohr and L. Selk. Estimating change points in nonparametric time series regression models. *Statistical Papers*, 2020;1-27.

[10] M. Kejriwal and P. Perron. A note on estimating a structural change in persistence. *Economics Letters*, 2012;117(3):932-935.