China's Regional Convergence in Panels with Multiple Structural Breaks

Matsuki, Takashi and Usami, Ryoichi

17 March 2007

China's Regional Convergence in Panels with Multiple Structural Breaks

(Running Title: China's Regional Convergence in Panels with Multiple Breaks)

Takashi Matsuki [*] and Ryoichi Usami

Department of Economics, Osaka Gakuin University,

2-36-1 Kishibeminami, Suita- City, Osaka, 564-8511, Japan

Abstract

This study investigates the existence of regional convergence of per capita outputs in China from 1952–2004, particularly focusing on considering the presence of multiple structural breaks in the provincial-level panel data. First, the panel-based unit root test that allows for occurrence of multiple breaks at various break dates across provinces is developed; this test is based on the p-value combination approach suggested by Fisher (1932). Next, the test is applied to China's provincial real per capita outputs to examine the regional convergence in China. To obtain the p-values of unit root tests for each province, which are combined to construct the panel unit root test, this study assumes three data generating processes: a driftless random walk process, an ARMA process, and an AR process with cross-sectionally dependent errors in Monte Carlo simulation. The results obtained from this study reveal that the convergence of the provincial per capita outputs exists in each of the three geographically classified regions—the Eastern, Central, and Western regions—of China.

[*] Corresponding author. E-mail: matsuki@ogu.ac.jp

1. Introduction

One of the important issues in China, which has achieved high economic growth rates since the end of 1978, is the existence of large differentials in output per capita between provinces. Reducing these gaps is one of the main objectives set in the Eleventh Five-Year-Plan (2006–2010). Therefore, from the perspective of policy making by the Chinese government, it is extremely important to understand the behaviour of provincial per capita outputs, particularly observing whether these per capita outputs can converge.

Lots of studies, including Mankiw, Romer and Weil (1992), Bernard and Durlauf (1995, 1996), and Quah (1993a, b, 1996), have been conducted on the convergence of per capita outputs since Barro (1991) and Barro and Sala-i-Martin (1992). Among these, some empirical studies have utilized nonstationary time series techniques such as unit root tests and cointegration tests.[1] On the other hand, Evans and Karras (1996), Lee, Pesaran and Smith (1997), Evans (1998), Flessig and Strauss (2001), and McCoskey (2002) have used unit root tests extended for panel data sets to investigate convergence across countries and the states of US; some of these tests have been proposed by Im, Pesaran and Shin (2003) (hereafter, IPS) and Maddala and Wu (1999) (hereafter, MW).

With regard to the convergence hypothesis of provincial per capita outputs in China, there are several contributions to the literature, such as Zhang, Liu and Yao (2001) and Pedroni and Yao (2006), that use unit root testing methods for a single time

---

[1] Bernard and Durlauf (1995), Oxley and Greasley (1995), Hobijn and Franses (2000), Pesaran (2004), Lim and McAleer (2004), etc.

series and panel data sets.[2] Zhang *et al.* (2001) aggregated the real per capita GDP of 30 Chinese provinces from 1952–1997 into three regions (the Eastern, Central, and Western regions) in accordance with the official classification and then applied the Augmented Dickey–Fuller t-test and the unit root test with a break suggested by Perron (1994) to the relative regional and national per capita GDPs of each of the three regions. Then, they concluded that two of the three regions (the Eastern and Western regions) are converging to their own specific steady states. Pedroni and Yao (2006) utilized the panel-based unit root tests, the IPS and MW tests, to investigate convergence of the annual real per capita GDP across all the provinces of China. They split each provincial series into the pre-reform sample (1952–1977) and the post-reform sample (1978–1997) to consider the impacts of the economic reforms since 1978. The results revealed convergence in the pre-reform sample, but not in the post-reform sample.

While testing unit roots or cointegrating relationships using a single time series, the sample size used in the analysis needs to be sufficiently large to obtain higher power of the test. Similarly, the time series dimension of panel sets for each cross-sectional unit should be large in panel unit root tests, especially in the tests based on combinations of separate unit root tests such as the IPS and MW tests. However, panels with longer time spans have a higher possibility of including structural changes caused by wars, supply shocks, significant policy changes, and so on. Perron (1989), Leybourne, Mills and

[2] The studies on regional growth in China which have not adopted the nonstationary time series or panel techniques are Chen and Fleisher (1996), Jian, Sachs and Warner (1996), Gundlach (1997), Raiser (1998), and Weeks and Yao (2003).

Newbold (1998), and Im, Lee and Tieslau (2005) have shown that ignoring the existing structural breaks in time series or panel data sets may lead to a substantial loss of power or serious size distortions in commonly used unit root tests such as the Dickey–Fuller test and the IPS test. Taking such evidence into account, it is desirable to employ tests that allow for structural changes in data.[3]

Smyth and Inder (2004) ascribe the logic behind the inclusion of multiple structural breaks in the official output of China to the occurrence of significant political and economic events: the Great Leap Forward from 1958–1960; the sudden suspension of the economic assistance from the Soviet Union in 1960; the crop failures from 1959–1961; the Cultural Revolution from 1966–1976; economic reforms from 1978–1979; and Deng Xiaoping's southern tour in 1992. Figures 1 to 3 show the fluctuations of log of real per capita outputs of the provinces, wherein each series is subtracted from the mean value of each of three regions, with twenty-nine provinces.[4] [5] In all the figures, we observe apparent shifts in the level of the series corresponding to

[3] Carlino and Mills (1993), Greasley and Oxley (1997), and Li and Papell (1999) have examined convergence using unit root tests which can deal with a breaking time series.

[4] This classification of provinces is nearly identical to that of Zhang *et al.* (2001), but the aggregation of provincial series is not conducted in this paper. The details will be described in Section 4.1.

[5] Studies on multi-country convergence often use the deviation from the cross-sectional mean and look into its nonstationarity (Evans and Karras, 1996; Lee *et al.*, 1997; and Evans, 1998), which will be described in the later sections.

each province for the following time periods: 1959–61, 1967–71, and the early 1990s (shown as grey areas in the figures). These shifts coincide with the occurrences of the events mentioned above. Based on these findings, some studies have focused on the presence of structural changes in the annual series in China (Li, 2000; Zhang *et al.*, 2001; and Smyth and Inder, 2004); these studies have adopted unit root testing methods permitting one or two breaks in a single time series for the analysis of the nonstationarity of the macroeconomic or provincial time series.

However, with regard to the convergence hypothesis in regional panel data in China, published papers which explicitly deal with the existence of multiple structural breaks occurring at different break dates in the panels have been few in number.[6]

---

[6] In general, existing panel-based unit root tests which allow for breaks may be too restrictive for empirical applications based on the convergence hypothesis. Specifically, these tests are based on two major assumptions: the presence of a linear time trend in a series and the absence of cross-sectional dependence between error terms in the data generating process (DGP). In the case of the former assumption, the tests defined under the DGP with a time trend are not directly applicable to investigations on convergence. In these investigations, the difference between two series or the deviation from the mean value of all cross-sectional units is usually used, and the difference or the mean deviation is often assumed to be zero mean stationary when absolute convergence exists, or level stationary when conditional convergence exists (Bernard and Durlauf, 1995; Evans and Karras, 1996). Thus, these analyses require tests which are defined under DGPs without a time trend, instead of DGPs with a time trend. In the case of the latter

Therefore, while examining convergence across provinces in China, this study focuses on the presence of multiple structural breaks in the panels. Based on the combining p-values method of Fisher (1932), we first develop a unit root test which allows for multiple breaks in the panels. We then investigate the existence of regional convergence in China by applying the test to China's provincial per capita outputs. The p-values of the t-type unit root tests for each province, which are combined by the panel unit root tests based on Fisher's p-values combination approach, are calculated by Monte Carlo simulation under three data generating models—the driftless random walk model, the ARMA model, and the AR model with cross-sectionally dependent errors. In particular, in the case of cross-sectional dependence between error terms, the bootstrap method proposed by Maddala and Wu (1999) and Wu and Wu (2001) is employed in order to correct the biases of the panel-based unit root tests.[7]

The remaining sections of this paper are organized as follows: Section 2 defines convergence; Section 3 describes the econometric methodology; Section 4 briefly mentions the data and discusses the empirical results; and Section 5 presents the conclusions.

---

assumption, cross-sectional correlation between error terms is a major issue in dynamic panel estimation because neglecting this correlation may lead to a bias of an estimated parameter and increase its variance (O'connell, 1998; Phillips and Sul, 2003).

[7] Banerjee and Carrion-i-Silvestre (2006) have dealt with several issues on structural breaks and cross-sectional dependence in a nonstationary panel framework.

2. Definition of Convergence

At first, we consider convergence as proposed by Evans (1998). Suppose that y_{it} is a log per capita output for province (cross-sectional unit) i at time t ($i = 1, \ldots, N$, $t = 1, \ldots, T$). Next, consider the difference between y_{it} and the mean value of y_{it} over $i = 1, \ldots, N$, which is denoted as $\tilde{y}_{it} \equiv y_{it} - \bar{y}_t$, where $\bar{y}_t \equiv N^{-1} \sum_{i=1}^{N} y_{it}$. As proved by Evans (1998), since $y_{it} - \bar{y}_t = N^{-1} \sum_{j=1}^{N} (y_{it} - y_{jt})$, if $y_{it} - y_{jt}$ is stationary for all pairs of i and j, $y_{it} - \bar{y}_t$ is stationary for all i. A converse proof is also available: since $y_{it} - y_{jt} = (y_{it} - \bar{y}_t) - (y_{jt} - \bar{y}_t)$, if $y_{it} - \bar{y}_t$ is stationary for all i, $y_{it} - y_{jt}$ is stationary for all pairs (i, j). These results equate bivariate convergence within all pairs of provinces, reflected by the stationarity of $y_{it} - y_{jt}$ for all pairs of i and j, to the stationarity of $y_{it} - \bar{y}_t$ for all i. This equivalence enables us to focus on investigating the stochastic properties of $\tilde{y}_{it} = y_{it} - \bar{y}_t$ for all i instead of $y_{it} - y_{jt}$ for all pairs of i and j.

In the next section, we will specify structural changes at some time periods in a series as multiple shifts in the level of the series. Accordingly, convergence is defined as follows:

For all i, if $\tilde{y}_{it}$ is stationary with shifts in its level at some t, then convergence exists across all the provinces.[8]

---

[8] Evans and Karras (1996) have postulated that convergence is absolute if $\tilde{y}_{it}$ has a zero mean for all i, or conditional if $\tilde{y}_{it}$ has a non-zero mean for some i. According to Evans and Karras, when all the series of $\tilde{y}_{it}$ are stationary and have some structural breaks, convergence can be considered as being absolute if $\tilde{y}_{it}$ has a zero mean for all

This study does not allow the trend stationarity of $\tilde{y}_{it}$ for each i because the presence of a linear time trend implies that some of the differences between y_{it} and y_{jt} for fixed i and all j will diverge as time approaches infinity unless the time trends are the same for all the pairs (Bernard and Durlauf, 1995). Further, with the exceptions of Liaoning in Figure 1 and Heilongjiang and Hubei in Figure 2, none of the figures show a distinct upward or downward tendency for any series during the entire sample period. Therefore, we consider $\tilde{y}_{it}$ as a series without a time trend in the later sections.

3. Econometric Methodology

3.1. Models and Test Statistics for a Single Time Series

We assume $\tilde{y}_{it}$ for each province to be nonstationary without breaks under the null hypothesis, and stationary with breaks under the alternative hypothesis. As discussed in Section 2, although each series exhibits no linear time trend, it contains some shifts in the level. Therefore, this study assumes that $\tilde{y}_{it}$ is generated by the following data generating process (DGP).

$$\text{\textit{Under Null}} \qquad \tilde{y}_{it} = \tilde{y}_{it-1} + \varepsilon_{it} \qquad\qquad (1)$$

$$\text{\textit{Under Alternative}} \qquad \tilde{y}_{it} = \rho_i \tilde{y}_{it-1} + \delta_{1i} DU_{1it} + \delta_{2i} DU_{2it} + \varepsilon_{it}, \ \ |\rho_i| < 1 \qquad (2)$$

$$i = 1, \ldots, N, \ \ t = 1, \ldots, T$$

---

i after the last break date, or conditional if $\tilde{y}_{it}$ has a non-zero mean for some i after the last break date.

where ε_{it} is independently and identically distributed across i and t with a zero mean and a finite variance; δ_{hi} denotes the size of the h th break ($h = 1, 2$); DU_{hit} denotes the h th break in the level of a series ($h = 1, 2$), where $DU_{hit} = 1$ for $t > \tau_{hi}T$, and zero otherwise; and τ_{hi} is the fraction of the h th break ($h = 1, 2$) in $0 < \tau_{1i} < \tau_{2i} < 1$, which is defined as TB_{hi} / T for all T, where TB_{hi} is the date of the h th break ($h = 1, 2$). In this DGP, the series is a driftless random walk process under the null hypothesis, whereas it is a stationary process and has up to two-time level shifts under the alternative hypothesis. Next, the regression model nests Models (1) and (2).

$$\Delta \widetilde{y}_{it} = \hat{\alpha}_{mi} d_m + \hat{\phi}_i \widetilde{y}_{it-1} + \hat{\delta}_{1i} DU_{1it} + \hat{\delta}_{2i} DU_{2it} + error \quad m = 1, 2 \tag{3}$$

where $\Delta \widetilde{y}_{it} = \widetilde{y}_{it} - \widetilde{y}_{it-1}$, $\hat{\phi}_i = \hat{\rho}_i - 1$, and d_m denotes the deterministic term, where $d_m = \{\varnothing\}$ for $m = 1$ and $\{1\}$ for $m = 2$. Let t_i^m be the t-statistic testing the null hypothesis $\hat{\phi}_i = 0$ and $\hat{\delta}_{1i} = \hat{\delta}_{2i} = 0$ against the alternative hypothesis $\hat{\phi}_i \neq 0$ and $\hat{\delta}_{1i} \neq 0$, $\hat{\delta}_{2i} \neq 0$ in each regression model m ($m = 1, 2$) for each i. As carried out in Zivot and Andrews (1992) and Lumsdane and Papell (1997), the break dates $\{TB_{1i}, TB_{2i}\}$ are endogenously determined to be where the one-sided t_i^m-statistic is minimized in sequential estimations over all possible break dates within the range of $0 < \tau_{1i} < \tau_{2i} < 1$. For fixed i, when Model (1) has a constant and Model (3) has both a constant and linear time trend, the t_i^m-statistic is the counterpart of the one proposed by Zivot and Andrews (1992) for a single break ($\delta_{2i} = 0$ and $\hat{\delta}_{2i} = 0$) and of that proposed by Lumsdane and Papell (1997) for double breaks, where the asymptotic behaviour of the statistic as $T \to \infty$ can be found. On the other hand, no literature provides the exact asymptotic behaviour of the test considered here. Therefore, as $T \to \infty$ for fixed i, we derive the limiting distribution of the statistic for each case of

breaks in the following theorems in which the subscript i is omitted for simplicity.

Theorem 1. *For Models (1) and (3), with $\delta_2 = 0$ and $\hat{\delta}_2 = 0$, as $T \to \infty$, the limiting distribution of the minimum t^m -statistic is given as follows:*

$$t^m(\tau_1) \Rightarrow \inf_{\tau_1} \left[\left\{ (1 + b_m^2) \int_0^1 W_m(r, \tau_1)^2 \, dr \right\}^{-1/2} \int_0^1 W_m(r, \tau_1) dW_m \right] \quad m = 1, \, 2 \qquad (4)$$

where $\Rightarrow$ denotes weak convergence in distribution; $W_m(r, \tau_1)$ denotes the residuals from the projection of a standard Wiener process $W(r)$ onto the subspace generated by the functions $\{du_1(r, \tau_1)\}$ for $m = 1$ and $\{1, du_1(r, \tau_1)\}$ for $m = 2$, where $du_1(r, \tau_1) = 1$ for $r > \tau_1$, and zero otherwise. b_m is given by

$$b_1 = (1 - \tau_1)^{-1} \left\{ \int_{\tau_1}^1 W(r) dr \right\}$$

$$b_2 = \tau_1^{-1} (1 - \tau_1)^{-1} \left\{ \int_{\tau_1}^1 W^\mu(r) dr \right\}$$

where $W^\mu(r)$ is a demeaned standard Wiener process defined as $W^\mu(r) \equiv W(r) - \int_0^1 W(r) dr$.

The proof of Theorem 1 is analogous to the following theorem and is, therefore, omitted.[9]

Theorem 2. *For Models (1) and (3), as $T \to \infty$, the limiting distribution of the minimum t_i^m -statistic is given as follows:*

$$t^m(\tau_1, \tau_2) \Rightarrow \inf_{\tau_1, \tau_2} \left[\left\{ (1 + b_m^2 + c_m^2) \int_0^1 W_m(r, \tau_1, \tau_2)^2 \, dr \right\}^{-1/2} \int_0^1 W_m(r, \tau_1, \tau_2) dW_m \right] \quad m = 1, \, 2 \quad (5)$$

---

[9] This proof is available on request.

where $W_m(r,\tau_1,\tau_2)$ denotes the residuals from the projection of $W(r)$ onto the subspace generated by the functions $\{du_1(r,\tau_1), du_2(r,\tau_2)\}$ for $m=1$ and $\{1, du_1(r,\tau_1), du_2(r,\tau_2)\}$ for $m=2$, where $du_h(r,\tau_h)=1$ for $r>\tau_h$, and zero otherwise ($h=1, 2$). b_m and c_m are given by

$$b_1 = (\tau_2 - \tau_1)^{-1}\left\{\int_{\tau_1}^1 W(r)dr - \int_{\tau_2}^1 W(r)dr\right\}$$

$$c_1 = (\tau_2 - \tau_1)^{-1}\left\{-\int_{\tau_1}^1 W(r)dr + (1-\tau_1)(1-\tau_2)^{-1}\int_{\tau_2}^1 W(r)dr\right\}$$

$$b_2 = (\tau_2 - \tau_1)^{-1}\left\{-\tau_1^{-1}\tau_2\int_0^{\tau_1} W^\mu(r)dr + \int_0^{\tau_2} W^\mu(r)dr\right\}$$

$$c_2 = (\tau_2 - \tau_1)^{-1}\left\{\int_0^{\tau_1} W^\mu(r)dr - (1-\tau_1)(1-\tau_2)^{-1}\int_0^{\tau_2} W^\mu(r)dr\right\}.$$

The proof of Theorem 2 is given in Appendix.

3.2. Construction of Panel Unit Root Test with Breaks

In this subsection, we construct a panel unit root test with breaks by combining the individual minimum t_i^m-test; this test is based on Fisher's (1932) sum of log p-value approach, which has been introduced and used by Maddala and Wu (1999).

Suppose that p_i is the p-value from the ith test statistic among N continuous test statistics. Therefore, since each p_i is an independent uniform ($0, 1$) variable, $-2\log p_i$ has the chi-square distribution with two degrees of freedom. Further, the summation of $-2\log p_i$ from $i=1$ to N also has the chi-square distribution with $2N$ degrees of freedom. Fisher (1932) utilized this fact to develop the test (hereafter, Fisher test). By applying Fisher's p-value combination method to N augmented Dickey–Fuller t-tests, Maddala and Wu (1999) has built a panel-based unit root test which does not allow breaks. In this study, we use Fisher's approach to construct panel unit root tests that allow multiple breaks. Let p_i^m denote the p-value from the

10

individual minimum t_i^m-test. Therefore, the sum of $\log p_i^m$ is defined as follows:

$$Fisher\_B = -2\sum_{i=1}^{N}\log p_i^m \qquad m = 1, 2 \qquad (6)$$

The Fisher_B test (the Fisher test with breaks) also has the chi-square distribution with $2N$ degrees of freedom. In the present case, however, the degree of freedom of the chi-square distribution is $2(N-1)$ due to the restriction of $\sum_{i=1}^{N}\widetilde{y}_{it} = 0$. The null and alternative hypotheses of the test are specified as H_0: $\phi_i = 0$ and $\delta_{1i} = \delta_{2i} = 0$ for all i and H_1: $\phi_i < 0$ and $\delta_{1i} \neq 0$, $\delta_{2i} \neq 0$ for some i respectively.

There are two noteworthy features of the tests based on Fisher's p-value combination approach: (1) Since the tests have an exact (chi-square) distribution, they do not require a large cross-sectional dimension of panel data. Hence, they are expected to perform well in the analyses using panels with a relatively large time dimension and a small cross-sectional dimension such as country-level, state-level, or provincial-level panels.[10] (2) Even if some of the N unit root tests give larger p-values than conventional significance levels, e.g. 5 or 10 per cent, which implies the non-rejection of the unit root null in each test, if these p-values indicate a slight tendency to reject the unit root null (e.g. 0.15 or 0.2), the tests based on Fisher's p-value combination approach can capture it.

To calculate the Fisher_B test statistic, we need to compute the p-value of the

_____

[10] Although Becker (1997) compared the performance of 16 p-value combination tests, including the Fisher test, he concluded that there was no test that was the most accurate or effective.

minimum t_i^m-test for all (i, m) by Monte Carlo simulation because the minimum

t_i^m-test has non-standard limiting distributions for each m shown in Theorems 1 and 2.

Under the unit root null hypothesis, this study considers the following three DGPs:

Model (I) $\qquad \widetilde{y}_{it} = \widetilde{y}_{it-1} + \varepsilon_{it}$

Model (II) $\qquad \hat{\theta}_i(L)\Delta\widetilde{y}_{it} = \hat{\psi}_i(L)\hat{\sigma}_i\varepsilon_{it}$

Model (III) $\qquad \Delta\widetilde{y}_{it}^* = \sum_{k=1}^{\bar{k}_i} \hat{\gamma}_{ik}\Delta\widetilde{y}_{it-k}^* + \varepsilon_{it}^*$

where ε_{it} is an $i.i.d.$ $N(0, 1)$ error across i and t;

$\hat{\theta}_i(L) = 1 - \hat{\theta}_{1i}L - \hat{\theta}_{2i}L^2 - \cdots - \hat{\theta}_{p_i}L^{p_i}$ and $\hat{\psi}_i(L) = 1 - \hat{\psi}_{1i}L - \hat{\psi}_{2i}L^2 - \cdots - \hat{\psi}_{q_i}L^{q_i}$, where

$\hat{\theta}_1, \cdots, \hat{\theta}_p$ and $\hat{\psi}_1, \cdots, \hat{\psi}_q$ are estimated parameters; and L is the lag operator such as

$Ly_t = y_{t-1}$. For Model (I), $\widetilde{y}_{it}$ is generated for each i by a driftless random walk

model. For Model (II), for each i, $\Delta\widetilde{y}_{it}$ is generated by the optimal autoregressive

moving average (ARMA) (p_i, q_i) model with estimated parameters and $N(0, \hat{\sigma}_i^2)$

innovations, where $\hat{\sigma}_i^2$ is the estimated innovation variance of the ARMA model. The

selection of the optimal ARMA model follows the Zivot and Andrews (1992) procedure,

which fits ARMA (p, q) model to $\Delta\widetilde{y}_t$ over the possible combinations of p and q

with $p, q \leq 5$, then finds the best fitted model according to the Akaike information

criterion and the Schwartz information criterion. When the two criteria choose different

models, the most parsimonious model is selected.

For Model (III), $\Delta\widetilde{y}_{it}^*$ is the bootstrap sample for $\Delta\widetilde{y}_{it}$, which is obtained by the

bootstrap method employed by Maddala and Wu (1999) and Wu and Wu (2001). The

procedure followed herein is elaborated below. Firstly, we estimate the equation

$\Delta\widetilde{y}_{it} = \sum_{k=1}^{\bar{k}_i} \hat{\gamma}_{ik}\Delta\widetilde{y}_{it-k} + \varepsilon_{it}^0$ for each i by using the OLS method, and then we obtain the

residuals $\varepsilon_t^0 = [\varepsilon_{1t}^0, \varepsilon_{2t}^0, \cdots, \varepsilon_{Nt}^0]$ ($t = 1,...,T$). Next, we resample ε_{it}^0 from the obtained

residuals by preserving their cross-sectional correlation structure based on the bootstrap

method of Maddala and Wu (1999), wherein the vector $\varepsilon_t^0 = [\varepsilon_{1t}^0, \cdots, \varepsilon_{Nt}^0]$ is resampled

instead of individual ε_{it}^0. In addition, we generate a random number g which takes

integer values on $[1, \ T]$ with probability $1/T$, by using a uniform random number.

We then draw a row of residuals $\varepsilon_g^0 = [\varepsilon_{1g}^0, \cdots, \varepsilon_{Ng}^0]$ according to the realizations of g.

The bootstrap sample ε_t^* ($t = 1,...,T$) is obtained by T-time withdrawals from the

residuals. The bootstrap sample $\tilde{y}_{it}^*$ is generated by Model (III) with estimated

parameters $\hat{\gamma}_{ik}$ ($k = 1,...,\bar{k}_i$) in the previous OLS estimation. However, $\tilde{y}_{i1}^*, \cdots, \tilde{y}_{i\bar{k}_i+1}^*$

are replaced by the sample obtained by the block resampling method of Berkowitz and

Kilian (1996). Their method divides the actual sample $\tilde{y}_{it}$ into $T - \bar{k}_i$ overlapping

subsampling blocks with size $\bar{k}_i + 1$ and randomly draws a block from $T - \bar{k}_i$ blocks.

Then, $\tilde{y}_{i1}^*, \cdots, \tilde{y}_{i\bar{k}_i+1}^*$ are replaced with this block.

In fact, in the case where the cross-sectionally dependent errors are present in the

data generating model, the Fisher_B test does not belong to the chi-square distribution

under the null hypothesis because the minimum t_i^m-tests are correlated across i.

Accordingly, the test may be biased towards over- or under-rejections of the null.

In order to correct these biases of the test, we first capture the cross-sectional

correlation structure in the panels according to the above resampling scheme.[11] Then,

with the generated bootstrap sample $\tilde{y}_{it}^*$ ($t = 1,...,T$), we obtain the empirical

---

[11] To remove cross-sectional dependence in the panels with structural breaks, the

common factor model is also applicable. See Banerjee and Carrion-i-Silvestre (2006).

distribution function of the Fisher_B test through simulation, which provides the appropriate small-sample critical values for the test. These values are listed in Table 2. Based on these simulated critical values, we can conduct unit root testing in an appropriate manner.

A Monte Carlo simulation is performed using 5,000 replications under each DGP. The summary of the simulation is as follows:

(1) For each i, the empirical distribution function of the minimum t_i^m-statistic is obtained through replications. In particular, in Model (III), 5,000 bootstrap samples are generated and used in the simulation.

(2) For each i, the p-value (p_i^m) of the actual minimum t_i^m-test, obtained from the original data set, is evaluated based on the empirical distribution function obtained in (1). Then, the Fisher_B statistic is calculated.

(3) In each replication in Model (III), p_i^m of the simulated minimum t_i^m-test, which is computed from each bootstrap sample, is evaluated for each i based on the empirical distribution function obtained in (1). Then, using $p_1^m, \ldots, p_N^m$, the value of $-2\sum_{i=1}^{N} \log p_i^m$ is calculated. The empirical distribution function of the Fisher_B test can be obtained from the calculated values of $-2\sum \log p_i^m$.

4. Empirical Analysis

4.1. Data

Provincial data have been sourced from *China Compendium of Statistics 1949–2004.* We have used the annual real per capita outputs of 29 provinces from 1952 to 2004;

14

these outputs have been generated by the chain index of the per capita gross regional product (GRP) with 1952 as the reference year.[12] Hainan and Sichuan provinces have been excluded due to the lack of data. All the series used in this paper have been taken in natural logarithms.[13]

As in Zhang *et al.* (2001), we divide the 29 provinces according to their geographical locations into the following three regions: the Eastern, Central and Western regions.[14] However, we have included the Guangxi Zhuang autonomous

---

[12] The chain index of the per capita GRP is computed as $Y_t^* = 100 \cdot (Y_{52}/100) \cdot (Y_{53}/100) \cdots (Y_t/100)$, where Y_t^* is the chain index of the per capita GRP. Further, Y_t is the index of the per capita GRP (preceding year = 100), and Y_{52} is set to 100.

[13] The quality of official Chinese statistics has been argued by many researchers (e.g. Chow, 1986; Rawski, 2001; and Holz, 2006). Currently, it is widely recognized that official Chinese data at the national and provincial levels have certain inconsistencies and miscalculations due to factors such as the lack of technical personnel for the collection of statistics and political pressure to exaggerate statistics at the lower levels. However, our results, which will be presented in Section 4.3, remain valid as long as the stochastic properties of the series used in this paper do not change even if there are certain inaccuracies in them.

[14] The Eastern region has the following ten provinces: Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Shandong, Fujian, and Guangdong. The Central region includes the following nine provinces: Shanxi, Inner Mongolia, Jilin,

15

region in the Western region, instead of the Eastern region, because since 1978, its log of real per capita output has shown considerable deviation from those of the other Eastern provinces. In fact, the differences between the recent data of Guangxi and other Western provinces are considerably less compared to the differences between Guangxi and other Eastern provinces. Therefore, it is reasonable to include Guangxi in the Western region.[15]

The panel for each region used in this study is composed of the deviations of a log of real per capita output from the mean value across all the provinces in the corresponding region, which is denoted by $\widetilde{y}_{it} = y_{it} - \sum_{i=1}^{N'} y_{it} = y_{it} - \bar{y}_t$, where N' is the number of provinces in the region.

4.2. Test Procedure

Model (3) shown in Section 3.1 is regressed for each m, including lagged augmentation terms of the first difference of $\widetilde{y}_{it}$, in order to eliminate the

---

Heilongjiang, Anhui, Jiangxi, Henan, Hubei, and Hunan. The Western region consists of the following ten provinces: Guangxi, Chongqing, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang.

[15] For example, for the series (in logarithms) in 2004, the difference between the series of Guangxi and Hebei (the closest series among other Eastern provinces) is 0.87. In contrast, the difference between the series of Guangxi and Yunnan (the closest among the Western provinces) is 0.09. In addition, the series of some other provinces in the Western region (Guizhou, Qinghai, and Xinjiang) are also close to that of Guangxi.

autocorrelation of the error term.

$$\Delta \tilde{y}_{it} = \hat{\alpha}_{mi} d_m + \hat{\phi}_i \tilde{y}_{it-1} + \hat{\delta}_{1i} DU_{1it} + \hat{\delta}_{2i} DU_{2it} + \sum_{l=1}^{\bar{l}_i} \hat{a}_{il} \Delta \tilde{y}_{it-l} + u_{it} \tag{7}$$

where $\bar{l}_i$ is a lag order parameter and u_{it} is a serially uncorrelated error. We determine the number of lagged augmentation terms by following the 'general-to-specific' procedure described in Perron (1989) and suggested in Ng and Perron (1995). The maximum lag order is set at 8. Next, the procedure first estimates the regression model with $\bar{l}_i = 8$. If the last lag is significant at 10 per cent, where the critical value is an asymptotic normal value of 1.645 on the t-statistic, the procedure selects 8 as the optimal lag order; otherwise, it is eliminated from the regression model. The steps mentioned above are repeated until the last lag becomes significant. In the event of a single insignificant lag, the optimal lag order is set at 0.

For each i, the minimum t_i^m-test statistic is obtained by sequentially regressing Model m ($m = 1, 2$) over the possible break dates $\{TB_{1i}, TB_{2i}\}$ within $1 + \bar{l}_i < TB_{1i} < TB_{2i} < 53$ for two-time breaks and $\{TB_{1i}\}$ within $1 + \bar{l}_i < TB_{1i} < 53$ for a one-time break. Then, for each of the three regions, the Fisher_B test is constructed for each m ($m = 1, 2$) by combining the p-value of the individual test (p_i^m), which is obtained via simulation.

4.3. Test Results and Discussion

We first employ the commonly used panel unit root tests without a break—the Levin, Lin and Chu (2002) test and the Im *et al.* (2003) test. The results are shown in Table 1. For each region, both the tests can reject the unit root null hypothesis in at least one

regression model at the 10 per cent or better significance level. From this test result, the convergence hypothesis of the provincial outputs appears to be supported for each region. However, the IPS and LLC tests may possibly suffer from biases towards under- or over-rejections of the unit root null because they do not treat the presence of both structural breaks and cross-sectional dependence among error terms in the panels.[16]

Next, we apply the tests based on Fisher's p-value combination approach—the MW test and the Fisher_B test—on series with breaks (The estimation results for each province in the presence of breaks are presented in Tables 1A–4A in Appendix.).[17] Table 2 provides the small-sample critical values at the 10, 5, and 1 per cent levels of the MW and Fisher_B tests under Model (III), which are obtained by using the procedure described in Section 3.2.

Table 3 reports the test results obtained under the three DGPs. In the case of tests on series without a break (the MW test), there are ten significant tests of regional

---

[16] With regard to these issues, Perron (1989), Leybourne *et al.* (1998), and Im *et al.* (2005) have revealed that ignoring breaks in a single time series or panel data can lead to an erroneous inference in a test, while O'connell (1998) and Phillips and Sul (2003) have argued that estimated parameters tend to be biased by the presence of cross-sectionally correlated errors.

[17] We have also obtained test results for cases in which the mean deviations of log per capita outputs display a linear time trend for Liaoning in the Eastern region and Heilongjiang and Hubei in the Central region. Since these results are quite similar to those tabulated in Table 3, they have not been reported but are available on request.

convergence of real per capita outputs. In these tests, however, due to the omission of breaks, the test results might be inaccurate and, therefore, misleading.

We then consider the possibility of structural breaks occurring at various break dates across provinces. The fourth column of Table 3 shows the results of the Fisher_B test in the case of a one-time break. When Models (I) and (II) are used as DGPs, for the Western region, the Fisher_B test rejects the unit root null hypothesis for both the regression models ($m = 1,\ 2$) at the 1 per cent significance level. In addition, under both the DGPs, significant rejections of the null are observed at the 10 per cent level for the Eastern region ($m = 1$) and at the 10 or 5 per cent level for the Central region ($m = 2$). In the case of Model (III), wherein there is the cross-sectional correlation between error terms, the test statistics for both the regression models for the Western region are still higher than the corresponding critical values at the 1 per cent significance level. Further, the statistic of the regression model for the Eastern region where $m = 1$ is also significant at the 10 per cent level. In the case of Central provinces, the Fisher_B test cannot support the stationarity alternative. In Model (III), the finding that convergence occurs within all provinces in the Eastern and Western regions appears to be consistent with that of Zhang *et al.* (2001).

The last column of Table 3 presents the results for cases with two-time breaks. In Models (I) and (II), with one exception in the Central region, all the test results for all the regions exhibit significant rejections of the unit root null hypothesis at the 5 per cent or better levels. Moreover, when the correlation of error terms among provinces in each region is considered in Model (III), the Fisher_B test also strongly supports the stationarity alternative with two-time shifts for all of the regions (for either or both of

19

the regression models). As compared to the case of a series that includes a single structural break, under any DGP, this case indicates the existence of regional convergence within all the three regions. Therefore, it should be concluded that dealing with multiple structural breaks occurring at different break dates for each province provides stronger evidence of the existence of convergence within regions in China. This fact may also account for the discrepancies in the results compared with those of Zhang *et al.* (2001), where one endogenous break point is assumed in their estimation.

The comparison of the three tests results shown in Table 3 reveals that they greatly depend on the number of breaks allowed in the tests. As discussed in Section 1, due to the impact of certain significant political and economic events, the provincial real per capita outputs in China are suspected to have some structural breaks; therefore, in the analysis on regional convergence in China, we consider it appropriate to examine the possibility of multiple structural changes in the studied time periods. Consequently, when the provincial log per capita outputs are allowed to have two-time level shifts at various break dates across the provinces, we observe convergence of the series in all the three regions.

4.4. Test Results Based on Other Regional Classifications [18]

As illustrated in Figure 1, the mean deviation of the real per capita output for Shanghai is much larger than those for other Eastern provinces. Since this may be indicative of the heterogeneity of Shanghai, the series for nine Eastern provinces,

---

[18] All the test results discussed in this subsection have been omitted but are available on request.

excluding Shanghai, have been tested. Consequently, convergence is also observed in the Eastern region.

Further investigations have been conducted based on other data classifications where the Eastern region (with or without Shanghai) includes the neighboring provinces, which are Guangxi, Jilin, and Heilongjiang. The cases where one, two, or all of the provinces are classified as belonging to the Eastern region are analyzed. As a result, in the case of two structural breaks, the evidence of convergence has been found in all the classifications. This fact seems to imply that the neighboring provinces are on the same path of convergence as that of other Eastern provinces; however, this is not conclusive.[19]

To make the discussion more concrete, in classifying provinces into certain regions, the use of classification methods such as cluster analysis would be desirable. The work of Hobijn and Franses (2000) is one such application. However, this is beyond the scope of this paper. Meanwhile, as discussed in Section 4.1, there appear to be substantial grounds for our classification of Chinese provinces. Therefore, our findings obtained from Table 3 are meaningful.

---

[19] In addition, the sample consisting of whole provinces has been tested; moreover, a significant rejection of the unit root null hypothesis has been obtained. However, we believe that further information (e.g. the homogeneity of provinces classified into different regions) is needed to arrive at a conclusion.

5. Conclusion

In this study, we investigated the regional convergence hypothesis of the provincial per capita outputs in China while considering up to two-time structural breaks in the panels. According to the p-value combination approach of Fisher (1932), the panel-based unit root test has been developed by combining the p-value of the individual unit root test which allows for breaks in a single time series. This approach allowed us to consider multiple breaks at various break dates across the provinces. We used three data generating models in the Monte Carlo simulation—the driftless random walk model, the ARMA model, and the AR model with cross-sectionally dependent errors—to calculate the p-value of the individual minimum t-type unit root test from its empirical distribution. In particular, in the case of the AR model with cross-sectionally dependent errors, the empirical distribution of the test for each province was generated on the bootstrap samples, which were obtained by the resampling procedure proposed by Maddala and Wu (1999) and Wu and Wu (2001). On the basis of their geographical locations, the provinces were grouped into the following three regions: the Eastern, Central, and Western regions. Subsequently, the existence of convergence within each region was tested by the panel unit root test with breaks, which was developed in this paper. As a result, when the presence of two-time breaks was considered in the test, we found significant evidence to suggest that the convergence of the provincial per capita outputs exists within each region.

participants at Yokohama Symposia in 2007 for their helpful comments.

References

Banerjee, A. and Carrion-i-Silvestre, J. L. (2006) Cointegration in panel data with breaks and cross-section dependence, *EUI Working Papers*, ECO No. 2006/5, 1-47.

Barro, R. J. (1991) Economic growth in a cross section of countries, *The Quarterly Journal of Economics*, **106**, 407-443.

Barro, R. J. and Sala-i-Martin, X. (1992) Convergence, *Journal of Political Economy*, **100**, 223-251.

Becker, B. J. (1997) P-values, combination of, in *Encyclopedia of Statistical Sciences: Update* vol.1 (Ed.) S. Kotz and N. L. Johnson, Wiley, New York, pp.448-453.

Berkowitz, J. and Kilian, L. (1996) Recent developments in bootstrapping time series, Discussion paper 96/54, Board of Governors of the Federal Reserve System.

Bernard, A. B. and Durlauf, S. N. (1995) Convergence in international output, *Journal of Applied Econometrics*, **10**, 97-108.

Bernard, A. B. and Durlauf, S. N. (1996) Interpreting tests of the convergence hypothesis, *Journal of Econometrics*, **71**, 161-173.

Carlino, G. A. and Mills, L. O. (1993) Are U.S. regional incomes converging? A time series analysis, *Journal of Monetary Economics*, **32**, 335-346.

Chen, J. and Fleisher, B. M. (1996) Regional income inequality and economic growth in China, *Journal of Comparative Economics*, **22**, 141-164.

Chow, G. C. (1986) Chinese statistics, *The American Statistician*, **40**, 191-196.

Evans, P. (1998) Using panel data to evaluate growth theories, *International Economic*

Review, **39**, 295-306.

Evans, P. and Karras, G. (1996) Convergence revisited, *Journal of Monetary Economics*, **37**, 249-265.

Fisher, R. A. (1932) *Statistical Methods for Research Workers*, Oliver & Boyd, Edinburgh, 4th Edition.

Fleissig, A. and Strauss, J. (2001) Panel unit-root tests of OECD stochastic convergence, *Review of International Economics*, **9**, 153-162.

Greasley, D. and Oxley, L. (1997) Time-series based tests of the convergence hypothesis: some positive results, *Economics Letters*, **56**, 143-147.

Gundlach, E. (1997) Regional convergence of output per worker in China: a neoclassical interpretation, *Asian Economic Journal*, **11**, 423-442.

Hobijn, B. and Franses, P. H. (2000) Asymptotically perfect and relative convergence of productivity, *Journal of Applied Econometrics*, **15**, 59-81.

Holz, C. A. (2006) Why China's new GDP data matters, *Far Eastern Economic Review*, **169**, 54-57.

Im, K.-S., Lee, J. and Tieslau, M. (2005) Panel LM unit-root tests with level shifts, *Oxford Bulletin of Economics and Statistics*, **67**, 393-419.

Im, K.-S., Pesaran, H. and Shin, Y. (2003) Testing for unit roots in heterogeneous panels, *Journal of Econometrics*, **115**, 53-74.

Jian, T., Sachs, J. D. and Warner, A. M. (1996) Trends in regional inequality in China, *China Economic Review*, **7**, 1-21.

Lee, K., Pesaran, M. H. and Smith, R. (1997) Growth and convergence in a multi-country empirical stochastic Solow model, *Journal of Applied Econometrics*,

12, 357-392.

Levin, A., Lin, C.-F. and Chu, C. (2002) Unit root test in panel data: Asymptotic and finite sample results, *Journal of Econometrics*, **108**, 1-24.

Leybourne, S. J., Mills, T. C. and Newbold, P. (1998) Spurious rejection by Dickey–Fuller tests in the presence of a break under the null, *Journal of Econometrics*, **87**, 191-203.

Li, Q. and Papell, D. (1999) Convergence of international output: Time series evidence for 16 OECD countries, *International Review of Economics and Finance*, **8**, 267-280.

Li, X.-M. (2000) The Great Leap Forward, economic reforms, and the unit root hypothesis: testing for breaking trend functions in China's GDP data, *Journal of Comparative Economics*, **28**, 814-827.

Lim, L. K. and McAleer, M. (2004) Convergence and catching up in ASEAN: a comparative analysis, *Applied Economics*, **36**, 137-153.

Lumsdaine, R. L. and Papell, D. H. (1997) Multiple trend breaks and the unit-root hypothesis, *The Review of Economics and Statistics*, **79**, 212-218.

Maddala, G. S. and Wu, S. (1999) A comparative study of unit root tests with panel data and a new simple test, *Oxford Bulletin of Economics and Statistics*, **61**, 631-652.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992) A contribution to the empirics of economic growth, *The Quarterly Journal of Economics*, **107**, 407-437.

McCoskey, S. K. (2002) Convergence in Sub-Saharan Africa: a nonstationary panel data approach, *Applied Economics*, **34**, 819-829.

National Bureau of Statistics (2005) *China Compendium of Statistics 1949–2004*, China Statistics Press, Beijing.

Ng, S. and Perron, P. (1995) Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag, *Journal of American Statistical Association*, **90**, 268-281.

O'connell, P. G. J. (1998) The overvaluation of purchasing power parity, *Journal of International Economics*, **44**, 1-19.

Oxley, L. and Greasley, D. (1995) A time-series perspective on convergence: Australia, UK and USA since 1870, *The Economic Record*, **71**, 259-270.

Pedroni, P. and Yao, J. Y. (2006) Regional income divergence in China, *Journal of Asian Economics*, **17**, 294-315.

Perron, P. (1989) The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, **57**, 1361-1401.

Perron, P. (1994) Trend, unit root and structural change in macroeconomic time series, in *Cointegration for the applied economists*, (Ed.) B. B. Rao, Macmillan, Basingstoke pp.113-146.

Pesaran, M. H. (2004) A pair-wise approach to testing for output and growth convergence, Mimeo, University of Cambridge.

Phillips, P. C. B. and Sul, D. (2003) Dynamic panel estimation and homogeneity testing under cross section dependence, *Econometrics Journal*, **6**, 217-259.

Quah, D. (1993a) Empirical cross-section dynamics in economic growth, *European Economic Review*, **37**, 426-434.

Quah, D. (1993b) Galton's fallacy and tests of the convergence hypothesis, *The Scandinavian Journal of Economics*, **95**, 427-443.

Quah, D. (1996) Empirics for economic growth and convergence, *European Economic*

Review, **40**, 1353-1375.

Raiser, M. (1998) Subsidising inequality: Economic reforms, fiscal transfers and convergence across Chinese provinces, *Journal of Development Studies*, **34**, 1-26.

Rawski, T. G. (2001) What is happening to China's GDP statistics? *China Economic Review*, **12**, 347-354.

Smyth, R. and Inder, B. (2004) Is Chinese provincial real GDP per capita nonstationary? Evidence from multiple trend break unit, *China Economic Review*, **15**, 1-24.

Weeks, M. and Yao, J. Y. (2003) Provincial conditional income convergence in China, 1953-1997: a panel data approach, *Econometric Reviews*, **22**, 59-77.

Wu, J.-L. and Wu, S. (2001) Is purchasing power parity overvalued? *Journal of Money, Credit, and Banking*, **33**, 804-812.

Zhang, Z., Liu, A. and Yao, S. (2001) Convergence of China's regional incomes 1952-1997, *China Economic Review*, **12**, 243-258.

Zivot, E. and Andrews, D. W. K. (1992) Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis *Journal of Business & Economic Statistics*, **10**, 251-270.

Figure 1. *Mean deviations of log of provincial real per capita outputs for the Eastern provinces*
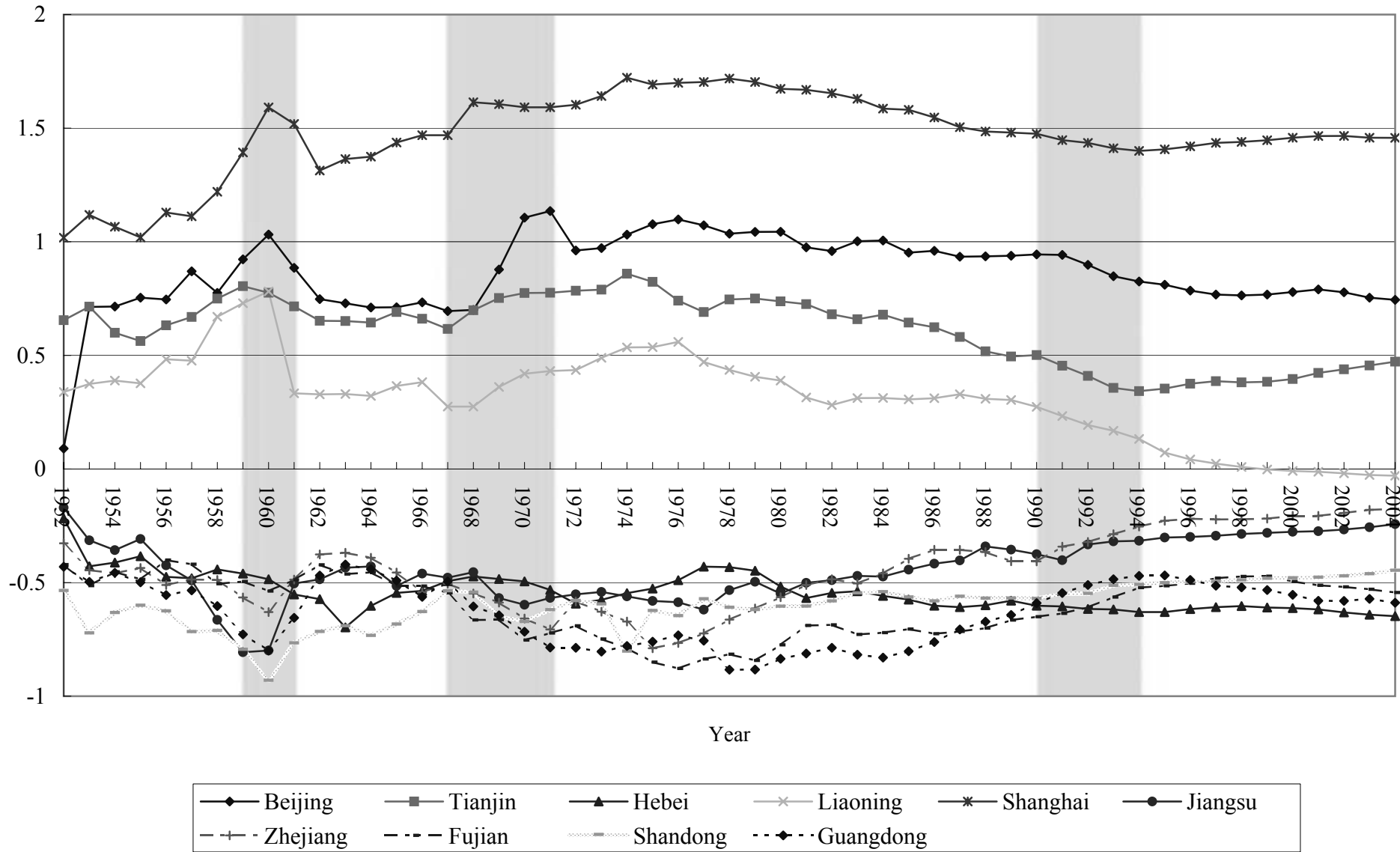
Figure 2. *Mean deviations of log of provincial real per capita outputs for the Central provinces*

Year

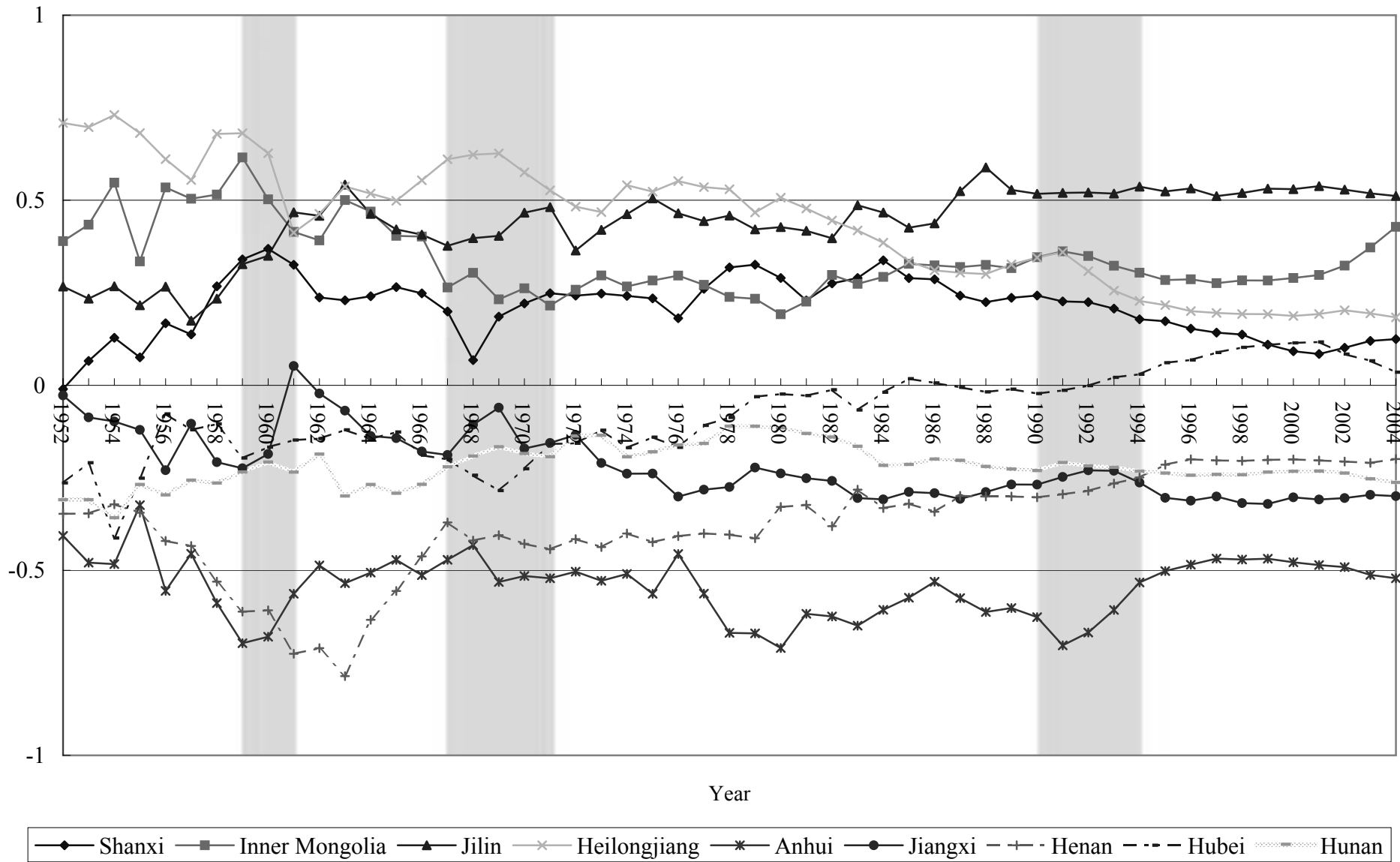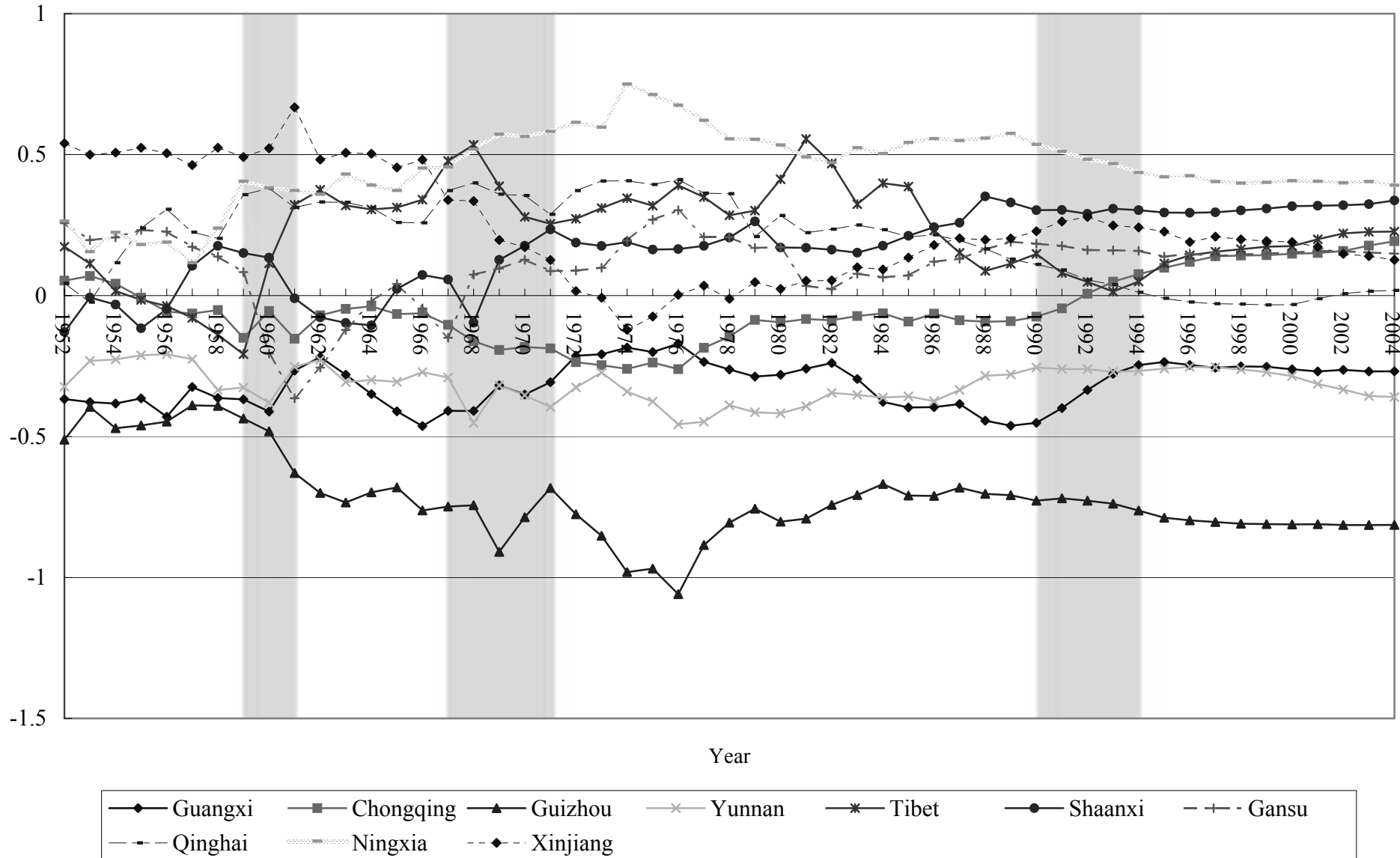Shanxi — Inner Mongolia — Jilin — Heilongjiang — Anhui — Jiangxi — Henan — Hubei — Hunan

Figure 3. *Mean deviations of log of provincial real per capita outputs for the Western provinces*

Table 1. *The results for the Levin et al. (2002) (LLC) test and the Im et al. (2003) (IPS) test*

Region	Regression Model	LLC test	IPS test[a]
East	no constant & no trend	–2.133**	–2.046**
	constant	2.564	0.816
Central	no constant & no trend	–1.669**	–0.569
	constant	–0.840	–1.366*
West	no constant & no trend	–1.947**	–1.648**
	constant	–1.486*	–1.893**

***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

[a]For both regression models, the means and the variances of the individual augmented Dickey–Fuller t-test for $T = 53 - p_i - 1$ were computed with 500,000 replications, where p_i is the number of lagged augmentation terms of the first difference of a series added in the individual ADF equation.

Table 2. *The critical values of the Maddala and Wu (1999) test and the Fisher_B test in the case of cross-sectionally dependent errors*

Test	Region	Regression Model[a]	10%	5%	1%[b]
MW test	East	$m = 1$	29.520	32.913	40.346
		$= 2$	29.428	32.926	38.824
	Central	$m = 1$	25.955	28.588	34.427
		$= 2$	26.242	28.986	35.202
	West	$m = 1$	28.205	31.419	37.292
		$= 2$	28.480	31.332	36.758
Fisher_B test					
one break	East	$m = 1$	30.102	33.585	41.364
		$= 2$	28.948	32.214	38.420
	Central	$m = 1$	25.972	28.672	35.302
		$= 2$	25.990	29.111	34.058
	West	$m = 1$	28.692	31.502	37.736
		$= 2$	28.575	31.364	36.900
two breaks	East	$m = 1$	29.950	33.425	40.936
		$= 2$	29.044	32.473	38.401
	Central	$m = 1$	26.181	29.165	35.346
		$= 2$	26.117	29.173	34.593
	West	$m = 1$	28.394	31.736	38.117
		$= 2$	28.709	31.942	38.513

[a]The regression model is $\Delta \tilde{y}_{it} = \hat{\alpha}_{mi} d_m + \hat{\phi}_i \tilde{y}_{it-1} + \hat{\delta}_{1i} DU_{1it} + \hat{\delta}_{2i} DU_{2it} + \sum_{l=1}^{\bar{l}_i} \hat{a}_{il} \Delta \tilde{y}_{it-l} + u_{it}$, $i = 1, \ldots, N'$, $t = 2, \ldots, 53$, where $N' = 10$ for the Eastern and Western regions and $N' = 9$ for the Central region, and $d_m = \{\emptyset\}$ for $m = 1$ and $d_m = \{1\}$ for $m = 2$; in addition, $\hat{\delta}_{1i} = \hat{\delta}_{2i} = 0$ for all i for the MW test and $\hat{\delta}_{2i} = 0$ for all i for the Fisher_B test for the one break case.

[b]The values are 10, 5, and 1 per cent points on the right tail of the empirical distributions of the MW and Fisher_B tests. These distributions are obtained as follows. For each i and m, under Model (III), the empirical distribution of the minimum t_i^m-statistic is obtained by a Monte Carlo simulation with 5,000 replications. Next, the percentage point (p_i^m) of the minimum t_i^m-statistic computed for each replication is evaluated on the empirical distribution obtained in the first step. After this, the value of $-2 \sum_{i=1}^{N} \log p_i^m$ is calculated for each replication. The empirical distributions of the tests can thus be obtained from the calculated values of $-2 \sum \log p_i^m$.

Table 3. *The results for the Maddala and Wu (1999) test and the Fisher_B test in the cases of one-time and two-time breaks*

Region	Regression Model[a]	MW test[b] (no break)	Fisher_B test[b] one break	Fisher_B test[b] two breaks
DGP Model (I): $\tilde{y}_{it} = \tilde{y}_{it-1} + \epsilon_{it}$				
East	$m = 1$	37.511***	27.251#[c]*	39.156#***
	$= 2$	12.125	16.436	33.962**
Central	$m = 1$	19.276	13.907	28.053**
	$= 2$	25.340*	28.723**	41.961***
West	$m = 1$	26.270*	37.805***	52.676#***
	$= 2$	26.751*	50.076#***	58.674#***
DGP Model (II): $\hat{\theta}_i(L)\Delta\tilde{y}_{it} = \hat{\psi}_i(L)\hat{\sigma}_i\epsilon_{it}$				
East	$m = 1$	35.940***	26.638#*	37.975***
	$= 2$	12.250	16.024	31.379**
Central	$m = 1$	17.185	9.875	19.740
	$= 2$	23.236	23.881*	36.547***
West	$m = 1$	26.214*	38.109#***	49.923***
	$= 2$	27.492*	49.960#***	58.591#***
DGP Model (III): $\Delta\tilde{y}_{it}^* = \sum_{k=1}^{\bar{k}_i} \hat{\gamma}_{ik}\Delta\tilde{y}_{it-k}^* + \epsilon_{it}^*$				
East	$m = 1$	19.601	31.744#*	53.414***
	$= 2$	15.104	18.074	22.887
Central	$m = 1$	17.042	23.211	36.035***
	$= 2$	27.109*	24.817	29.930**
West	$m = 1$	28.991*	41.450***	45.013***
	$= 2$	28.882*	37.688***	34.546**

***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

[a]The regression model is $\Delta\tilde{y}_{it} = \hat{\alpha}_{mi}d_m + \hat{\phi}_i\tilde{y}_{it-1} + \hat{\delta}_{1i}DU_{1it} + \hat{\delta}_{2i}DU_{2it} + \sum_{l=1}^{\bar{l}_i} \hat{a}_{il}\Delta\tilde{y}_{it-l} + u_{it}$, $i = 1,\ldots,N'$, $t = 2,\ldots,53$, where $N' = 10$ for the Eastern and Western regions and $N' = 9$ for the Central region, and $d_m = \{\emptyset\}$ for $m = 1$ and $d_m = \{1\}$ for $m = 2$; in addition, $\hat{\delta}_{1i} = \hat{\delta}_{2i} = 0$ for all i for the MW test and $\hat{\delta}_{2i} = 0$ for all i for the Fisher_B test for the one break case.

[b]Under Models (I) and (II), because of the restriction of $\sum_{i=1}^{N} \tilde{y}_{it} = 0$, the degree of freedom of the chi-square distribution of the test is $2(N - 1)$.

[c]The sign # indicates that the p-values for some provinces were estimated to be zero due to the fact that for each of these provinces, the realization of the minimum t_i^m-statistic lay far left from its empirical distribution which was generated by a Monte Carlo simulation with 5,000 replications. Therefore, in order to calculate the Fisher_B statistic, the obtained p-values for these provinces were assigned a value of 0.0002 (1/5000). This implies that we assume that the minimum t_i^m-statistic for each of these provinces took a value within the estimated empirical distribution only once in the 5,000 replications.

Appendix

Proof of Theorem

For simplicity, we omit the subscript i of a variable and denote a time series as merely y_t instead of $\tilde{y}_t$ used in the main text. Therefore, y_t is assumed to be subject to Models (1) and (2) with an *i.i.d.* innovation ε_t with a zero mean and a finite variance σ^2. In this proof, we show the derivation only for the case of $m = 2$ because that for the case of $m = 1$ is obtained along the same lines.

Let e_t be the OLS residual obtained by regressing y_t on an intercept and two dummy variables (DU_{1t} and DU_{2t}) for $t = 1, \ldots, T$. Then, the residual is expressed as

$$e_t = S_t^\mu - \hat{\delta}_1 \left(DU_{1t} - \overline{DU_1} \right) - \hat{\delta}_2 \left(DU_{2t} - \overline{DU_2} \right) \tag{1A}$$

where S_t^μ is the demeaned random walk process such as $S_t^\mu \equiv S_t - T^{-1} \sum_{t=1}^{T} S_t$, where $S_t = \sum_{s=1}^{t} \varepsilon_s$, and $\overline{DU_h} = \sum_{t=1}^{T} DU_{ht} = 1 - \tau_h$ ($h = 1, 2$). Now, we write $\hat{\boldsymbol{\delta}} = (\mathbf{X'X})^{-1} \mathbf{X'S}^\mu$, where $\hat{\boldsymbol{\delta}} = (\hat{\delta}_1, \hat{\delta}_2)'$; $\mathbf{X} = \left(\mathbf{DU_1} - \overline{\mathbf{DU_1}}, \mathbf{DU_2} - \overline{\mathbf{DU_2}} \right)$, where $\mathbf{DU_h} - \overline{\mathbf{DU_h}} = \left(DU_{h1} - \overline{DU_h}, \cdots, DU_{hT} - \overline{DU_h} \right)'$; $\mathbf{S}^\mu = \left(S_1^\mu, \cdots, S_T^\mu \right)'$. Then, we have

$$T^{-1/2} \hat{\boldsymbol{\delta}} = \frac{1}{(\tau_2 - \tau_1) T^{3/2}} \left(\begin{array}{c} -\tau_1^{-1} \tau_2 \sum_{t=1}^{\tau_1 T} S_t^\mu + \sum_{t=1}^{\tau_2 T} S_t^\mu \\ -\tau_1 \sum_{t=1}^{T} S_t^\mu + \sum_{t=1}^{\tau_1 T} S_t^\mu + (1 - \tau_2)^{-1} (1 - \tau_1) \left(\tau_2 \sum_{t=1}^{T} S_t^\mu - \sum_{t=1}^{\tau_2 T} S_t^\mu \right) \end{array} \right).$$

Thus,

$$T^{-1/2} \begin{pmatrix} \hat{\delta}_1 \\ \hat{\delta}_2 \end{pmatrix} \Rightarrow \frac{\sigma}{\tau_2 - \tau_1} \left(\begin{array}{c} -\tau_1^{-1} \tau_2 \int_0^{\tau_1} W^\mu(r) dr + \int_0^{\tau_2} W^\mu(r) dr \\ \int_0^{\tau_1} W^\mu(r) dr - (1 - \tau_2)^{-1} (1 - \tau_1) \int_0^{\tau_2} W^\mu(r) dr \end{array} \right) = \sigma \begin{pmatrix} A \\ B \end{pmatrix}$$

where $\Rightarrow$ denotes weak convergence in distribution. Therefore,

$$T^{-1/2} e_t = T^{-1/2} S_t^\mu - T^{-1/2} \hat{\delta}_1 \left(DU_{1t} - \overline{DU_1} \right) - T^{-1/2} \hat{\delta}_2 \left(DU_{2t} - \overline{DU_2} \right)$$

$$\Rightarrow \sigma W_2(r, \tau_1, \tau_2) \tag{2A}$$

$$= \sigma \left[W^{\mu}(r) - A\{du_1(r,\tau_1) - (1-\tau_1)\} - B\{du_2(r,\tau_2) - (1-\tau_2)\} \right]$$

where $W_2(r,\tau_1,\tau_2)$ denotes the residual from the projection of $W(r)$ onto the subspace generated by the function $\{1, du_1(r,\tau_1), du_2(r,\tau_2)\}$, where $du_h(r,\tau_h) = 1$ for $r > \tau_h$, and zero otherwise ($h = 1, 2$).

From the regression of e_t on e_{t-1}, we can obtain the t-test statistic as

$$t = \frac{T^{-1}\sum_{t=2}^{T} e_{t-1}\Delta e_t}{s\left(T^{-2}\sum_{t=2}^{T} e_{t-1}^2\right)^{1/2}} \tag{3A}$$

where $s^2 = T^{-1}\sum_{t=2}^{T}(e_t - \hat{\rho}e_{t-1})^2$, where $\hat{\rho}$ is the estimated coefficient of e_{t-1} in the regression of e_t on e_{t-1}. Now, we show the probability limits of $T^{-1}\sum_{t=2}^{T} e_{t-1}\Delta e_t$, $T^{-2}\sum_{t=2}^{T} e_{t-1}^2$ and s^2.

$T^{-1}\sum e_{t-1}\Delta e_t$ is expressed as

$$T^{-1}\sum_{t=2}^{T} e_{t-1}\Delta e_t = T^{-1}\sum_{t=2}^{T} S_{t-1}^{\mu}\Delta S_t^{\mu} - T^{-1}\hat{\delta}_1\left\{S_{\tau_1 T}^{\mu} + \tau_1 y_T - y_{\tau_1 T+1} + (1-\tau_1)y_1 + (1-\tau_1)\hat{\delta}_1\right\}$$

$$- T^{-1}\hat{\delta}_2\left\{S_{\tau_2 T}^{\mu} + \tau_2 y_T - y_{\tau_2 T+1} + (1-\tau_2)y_1 + (1-\tau_2)\hat{\delta}_2\right\} + (\tau_1 + \tau_2 - 1)T^{-1}\hat{\delta}_1\hat{\delta}_2.$$

Hence, we have the following limiting distribution.

$$T^{-1}\sum_{t=2}^{T} e_{t-1}\Delta e_t \Rightarrow \sigma^2\left[\int_0^1 W^{\mu}(r)dW^{\mu}(r) - A\left\{-\int_0^1 W(r)dr + \tau_1 W(1) + (1-\tau_1)A\right\}\right.$$

$$\left. - B\left\{-\int_0^1 W(r)dr + \tau_2 W(1) + (1-\tau_2)B\right\} + (\tau_1 + \tau_2 - 1)AB\right].$$

In the derivation above, we used the following facts that

$$T^{-1}\sum_{t=2}^{T} S_{t-1}^{\mu}\Delta S_t^{\mu} \Rightarrow \sigma^2\int_0^1 W^{\mu}(r)dW^{\mu}(r), \quad T^{-1/2}S_{\tau_h T}^{\mu} \Rightarrow \sigma W^{\mu}(\tau_h)$$

$$T^{-1/2}y_T \Rightarrow \sigma W(1), \quad T^{-1/2}y_{\tau_h T+1} \Rightarrow \sigma W(\tau_h), \quad T^{-1/2}y_1 \Rightarrow \sigma W(0).$$

Based on the limiting behaviour of $T^{-1/2}e_t$ shown in Equation (2A) and the continuous mapping theorem, it is straightforward to show that

$$T^{-2}\sum_{t=2}^{T}e_{t-1}^2 \Rightarrow \sigma^2\int_0^1 W_2(r,\tau_1,\tau_2)dr.$$

Finally, since $s^2 = T^{-1}\sum_{t=2}^{T}(e_t - \hat{\rho}e_{t-1})^2 = T^{-1}\sum_{t=2}^{T}\Delta e_t^2 + o_p(1)$, we show the probability limit of the first term in the last equation.

$$T^{-1}\sum_{t=2}^{T}\Delta e_t^2 = T^{-1}\sum_{t=2}^{T}\Delta y_t^2 + T^{-1}\left(\hat{\delta}_1^2 + \hat{\delta}_2^2\right) - 2T^{-1}\left\{\hat{\delta}_1(y_{\tau_1 T+1} - y_{\tau_1 T}) + \hat{\delta}_2(y_{\tau_2 T+1} - y_{\tau_2 T})\right\}$$

$$\Rightarrow \sigma^2\left(1 + A^2 + B^2\right).$$

Note that

$$T^{-1}\sum_{t=2}^{T}\Delta y_t^2 \xrightarrow{\;p\;} \sigma^2, \quad T^{-1/2}(y_{\tau_h T+1} - y_{\tau_h T}) \xrightarrow{\;p\;} 0$$

where $\xrightarrow{\;p\;}$ denotes convergence in probability. ∎

Estimation Results for Each Province

The estimation results, obtained from individual regression conducted for each province, are shown in the following tables. Now, we briefly discuss the estimation accuracy of the dates of breaks and signs of break size parameters in the results. Herein, the time periods of influential events in China, as described in Section 1, are considered as 1958–62, 1966–72, and 1989–95 (i.e. the time periods shown as grey areas in Figures 1–3 with a one-year lead and lag for each period). In the case of a one-time break (shown in Tables 1A and 2A), approximately one-third to half of the estimated break dates are consistent with the expected time periods, for each region. For the size parameters of the break variables that have the expected break dates, some have the right sign while others do not. In the case of two-time breaks (shown in Tables 3A and 4A), approximately half of the detected break dates (of the total number of the first and second breaks) match the expected periods, for each region. For the sign of break size parameters, there is some improvement in terms of accuracy.

Table 1A. *The estimation results for each province in the case of a one-time break (m = 1)*

Region	Province	$\hat{\phi}$	$\hat{\delta}_1$	Min t	$\bar{l}$	Break Date	P-value Model (I)	Model (II)	Model (III)
East	Beijing	−0.094	0.087	−2.893	6	1963	0.356	0.372	0.092
	Tianjin	−0.050	0.031	−2.110	6	1962	0.678	0.724	0.500
	Hebei	−0.002	−0.006	−0.025	0	1954	0.994	0.994	0.932
	Liaoning	−0.081	−0.047	−2.650	8	1994	0.459	0.464	0.494
	Shanghai	−0.149	0.227	−7.582	8	1962	0.000#[a]	0.000#	0.000#
	Jiangsu	−0.042	0.012	−2.838	8	2003	0.379	0.386	0.451
	Zhejiang	−0.046	−0.017	−1.122	1	1955	0.937	0.939	0.892
	Fujian	−0.050	−0.034	−1.716	5	1962	0.813	0.804	0.513
	Shandong	−0.036	0.008	−3.161	8	2003	0.261	0.303	0.378
	Guangdong	−0.092	−0.060	−1.986	1	1955	0.728	0.748	0.386
Central	Shanxi	−0.235	0.055	−3.839	8	1968	0.078	0.118	0.048
	Inner Mongolia	−0.071	0.033	−2.516	8	1980	0.516	0.580	0.218
	Jilin	−0.053	0.033	−0.592	6	1960	0.981	0.986	0.704
	Heilongjiang	−0.036	0.004	−2.377	3	1998	0.577	0.642	0.527
	Anhui	−0.121	−0.067	−1.522	0	1954	0.865	0.874	0.650
	Jiangxi	−0.287	−0.095	−3.491	8	1969	0.149	0.554	0.025
	Henan	−0.016	−0.003	−0.991	4	1996	0.950	0.951	0.945
	Hubei	−0.049	0.012	−1.248	4	1991	0.915	0.912	0.876
	Hunan	−0.090	−0.034	−2.844	8	1981	0.376	0.395	0.176
West	Guangxi	−0.057	−0.017	−1.528	0	1961	0.859	0.866	0.739
	Chongqing	−0.064	0.020	−1.556	2	1991	0.854	0.856	0.815
	Guizhou	−0.006	−0.014	−0.274	7	1983	0.989	0.991	0.898
	Yunnan	−0.270	−0.091	−2.832	8	1962	0.384	0.425	0.104
	Tibet	−0.253	0.072	−3.449	2	1959	0.170	0.183	0.148
	Shaanxi	−0.379	0.101	−4.550	0	1968	0.021	0.015	0.047
	Gansu	−1.229	0.175	−7.206	7	1972	0.000#	0.000#	0.0002
	Qinghai	−0.045	0.019	−1.863	8	1999	0.773	0.780	0.670
	Ningxia	−0.196	0.101	−3.891	0	1957	0.079	0.076	0.036
	Xinjiang	−0.089	0.020	−2.496	8	1982	0.518	0.518	0.540

[a]The sign # indicates that the p-value for the province was estimated to be zero due to the fact that the realization of the minimum t_i^m-statistic lay far left from its empirical distribution which was generated by a Monte Carlo simulation with 5,000 replications. Therefore, in order to calculate the Fisher_B statistic, the obtained p-value was assigned a value of 0.0002 (1/5000). This implies that we assume that the minimum t_i^m-statistic took a value within the estimated empirical distribution only once in the 5,000 replications.

Table 2A. *The estimation results for each province in the case of a one-time break (m = 2)*

Region	Province	$\hat{\phi}$	$\hat{\delta}_1$	Min t	$\bar{l}$	Break Date	P-value		
							Model (I)	Model (II)	Model (III)
East	Beijing	−0.275	0.064	−3.628	6	1967	0.385	0.410	0.555
	Tianjin	−0.256	−0.066	−3.429	6	1986	0.479	0.527	0.439
	Hebei	−0.481	−0.045	−3.814	5	1983	0.295	0.297	0.433
	Liaoning	−0.349	−0.112	−2.786	6	1990	0.779	0.765	0.466
	Shanghai	−0.295	−0.084	−4.700	6	1983	0.061	0.058	0.126
	Jiangsu	−0.294	0.079	−2.312	8	1985	0.908	0.896	0.567
	Zhejiang	−0.153	0.060	−2.722	0	1979	0.802	0.791	0.734
	Fujian	−0.144	−0.029	−2.768	5	1964	0.788	0.779	0.785
	Shandong	−0.452	0.070	−4.007	0	1976	0.228	0.264	0.071
	Guangdong	−0.158	0.026	−2.744	1	1986	0.795	0.799	0.830
Central	Shanxi	−0.582	−0.065	−4.340	5	1993	0.131	0.187	0.183
	Inner Mongolia	−0.314	0.048	−3.999	7	2000	0.230	0.276	0.324
	Jilin	−0.611	0.057	−5.479	6	1986	0.010	0.011	0.028
	Heilongjiang	−0.214	−0.066	−2.830	3	1980	0.743	0.817	0.640
	Anhui	−0.553	−0.025	−4.036	3	1976	0.216	0.217	0.220
	Jiangxi	−0.574	−0.079	−4.460	0	1974	0.104	0.450	0.133
	Henan	−0.286	0.065	−3.965	3	1979	0.239	0.284	0.279
	Hubei	−0.216	0.051	−2.921	4	1976	0.710	0.719	0.630
	Hunan	−0.273	0.018	−2.959	0	1965	0.692	0.719	0.744
West	Guangxi	−0.340	0.022	−3.179	4	1966	0.604	0.589	0.603
	Chongqing	−0.199	0.059	−3.282	2	1990	0.556	0.550	0.493
	Guizhou	−0.475	−0.064	−4.346	7	1971	0.124	0.126	0.116
	Yunnan	−0.339	−0.016	−3.156	0	1969	0.618	0.640	0.632
	Tibet	−0.595	−0.119	−7.793	7	1985	0.000#[a]	0.000#	0.002
	Shaanxi	−0.380	0.097	−4.518	0	1968	0.089	0.089	0.224
	Gansu	−1.196	0.188	−6.392	7	1973	0.001	0.001	0.018
	Qinghai	−0.272	−0.071	−2.805	8	1987	0.766	0.773	0.548
	Ningxia	−0.213	0.095	−3.628	0	1957	0.379	0.382	0.425
	Xinjiang	−0.351	−0.122	−4.566	7	1966	0.084	0.086	0.160

[a]The sign # indicates that the p-value for the province was estimated to be zero due to the fact that the realization of the minimum t_i^m-statistic lay far left from its empirical distribution which was generated by a Monte Carlo simulation with 5,000 replications. Therefore, in order to calculate the Fisher‗B statistic, the obtained p-value was assigned a value of 0.0002 (1/5000). This implies that we assume that the minimum t_i^m-statistic took a value within the estimated empirical distribution only once in the 5,000 replications.

Table 3A. *The estimation results for each province in the case of two-time breaks (m = 1)*

Region	Province	$\hat{\phi}$	$\hat{\delta}_1$	$\hat{\delta}_2$	Min t	$\bar{l}$	Break Date		P-value		
							1st break	2nd break	Model (I)	Model (II)	Model (III)
East	Beijing	−0.066	0.159	−0.109	−3.394	6	1966	1971	0.710	0.743	0.140
	Tianjin	−0.114	0.090	−0.044	−3.916	8	1963	1982	0.485	0.558	0.206
	Hebei	−0.176	−0.093	−0.018	−3.258	1	1955	1979	0.757	0.775	0.117
	Liaoning	−0.180	0.133	−0.155	−4.771	8	1969	1978	0.158	0.163	0.175
	Shanghai	−0.155	0.243	−0.014	−7.869	8	1962	1990	0.000#[a]	0.0002	0.0002
	Jiangsu	−0.106	−0.100	0.102	−5.425	8	1968	1977	0.048	0.046	0.028
	Zhejiang	−0.151	−0.143	0.145	−4.405	8	1965	1979	0.273	0.270	0.151
	Fujian	−0.075	−0.060	0.025	−2.439	6	1962	1987	0.945	0.940	0.527
	Shandong	−0.057	−0.206	0.194	−4.504	8	1973	1974	0.238	0.308	0.234
	Guangdong	−0.139	−0.233	0.144	−4.894	8	1967	1970	0.129	0.150	0.042
Central	Shanxi	−0.278	0.163	−0.101	−5.415	8	1968	1970	0.042	0.090	0.025
	Inner Mongolia	−0.289	0.174	−0.085	−5.937	7	1961	1964	0.013	0.022	0.008
	Jilin	−0.078	0.042	0.006	−0.710	6	1960	1985	0.999	1.000	0.807
	Heilongjiang	−0.267	0.129	−0.099	−5.766	8	1962	1983	0.020	0.076	0.022
	Anhui	−0.137	−0.276	0.201	−2.204	0	1955	1956	0.967	0.970	0.691
	Jiangxi	−0.472	−0.052	−0.088	−4.849	8	1962	1969	0.122	0.517	0.026
	Henan	−0.107	−0.127	0.106	−2.667	1	1956	1963	0.914	0.914	0.700
	Hubei	−0.171	−0.037	0.050	−3.045	6	1962	1974	0.825	0.830	0.722
	Hunan	−0.099	−0.022	−0.015	−2.929	8	1980	1983	0.854	0.884	0.468
West	Guangxi	−0.091	−0.115	0.092	−2.133	4	1964	1966	0.974	0.979	0.863
	Chongqing	−0.213	−0.032	0.060	−3.319	3	1964	1989	0.727	0.743	0.667
	Guizhou	−0.109	−0.105	0.018	−2.381	0	1958	1970	0.954	0.958	0.389
	Yunnan	−0.217	−0.103	0.033	−3.110	0	1956	1960	0.801	0.850	0.186
	Tibet	−0.518	0.186	−0.112	−7.848	1	1959	1985	0.000#	0.0002	0.002
	Shaanxi	−0.781	0.147	0.094	−6.575	2	1968	1987	0.003	0.003	0.034
	Gansu	−1.360	−0.042	0.235	−7.621	7	1965	1972	0.000#	0.0006	0.006
	Qinghai	−0.276	0.087	−0.064	−3.911	8	1964	1987	0.473	0.522	0.351
	Ningxia	−0.227	0.127	−0.038	−4.662	0	1957	1990	0.180	0.191	0.058
	Xinjiang	−0.129	0.117	−0.102	−3.582	3	1957	1961	0.617	0.636	0.627

[a]The sign # indicates that the p-value for the province was estimated to be zero due to the fact that the realization of the minimum t_i^m-statistic lay far left from its empirical distribution which was generated by a Monte Carlo simulation with 5,000 replications. Therefore, in order to calculate the Fisher_B statistic, the obtained p-value was assigned a value of 0.0002 (1/5000). This implies that we assume that the minimum t_i^m-statistic took a value within the estimated empirical distribution only once in the 5,000 replications.

Table 4A. *The estimation results for each province in the case of two-time breaks (m = 2)*

| Region | Province | $\hat{\phi}$ | $\hat{\delta}_1$ | $\hat{\delta}_2$ | Min t | $\bar{l}$ | Break Date | | P-value | | |
							1st break	2nd break	Model (I)	Model (II)	Model (III)
East	Beijing	−0.473	0.115	−0.083	−6.668	6	1967	1991	0.012	0.017	0.106
	Tianjin	−0.285	0.025	−0.081	−3.875	6	1967	1986	0.833	0.854	0.770
	Hebei	−1.301	0.079	−0.139	−6.271	8	1969	1985	0.026	0.038	0.221
	Liaoning	−0.289	−0.097	−0.087	−3.615	0	1960	1988	0.898	0.894	0.684
	Shanghai	−0.401	0.119	−0.085	−5.721	8	1967	1983	0.087	0.095	0.205
	Jiangsu	−0.288	−0.050	0.093	−2.749	8	1968	1984	0.986	0.988	0.862
	Zhejiang	−0.238	−0.136	0.195	−4.517	8	1967	1979	0.542	0.552	0.530
	Fujian	−0.280	−0.073	0.056	−4.723	5	1966	1987	0.436	0.454	0.462
	Shandong	−0.640	0.084	0.072	−5.327	0	1965	1992	0.180	0.231	0.106
	Guangdong	−0.527	−0.154	0.099	−6.045	8	1967	1990	0.047	0.058	0.189
Central	Shanxi	−0.501	0.087	−0.062	−4.728	1	1955	1993	0.432	0.539	0.492
	Inner Mongolia	−0.765	−0.150	0.049	−4.876	8	1966	1983	0.366	0.462	0.532
	Jilin	−1.044	0.129	0.091	−7.276	6	1960	1986	0.002	0.004	0.023
	Heilongjiang	−0.382	−0.074	−0.059	−3.704	4	1980	1991	0.865	0.932	0.792
	Anhui	−0.860	−0.080	0.099	−6.422	3	1976	1993	0.023	0.020	0.025
	Jiangxi	−0.657	−0.089	−0.027	−5.493	0	1972	1994	0.145	0.480	0.272
	Henan	−0.312	0.107	0.045	−5.694	3	1963	1986	0.099	0.138	0.262
	Hubei	−0.489	0.065	0.088	−4.278	0	1955	1977	0.650	0.666	0.617
	Hunan	−1.231	0.112	−0.119	−6.698	8	1971	1986	0.012	0.015	0.060
West	Guangxi	−0.563	0.054	−0.029	−3.966	7	1968	1982	0.777	0.780	0.804
	Chongqing	−0.232	0.028	0.054	−3.691	2	1977	1991	0.866	0.877	0.807
	Guizhou	−0.655	−0.207	0.188	−6.036	8	1973	1976	0.053	0.048	0.057
	Yunnan	−0.623	−0.066	0.058	−4.957	0	1967	1986	0.326	0.372	0.519
	Tibet	−0.665	0.053	−0.171	−8.212	7	1980	1985	0.000#[a]	0.0002	0.016
	Shaanxi	−0.784	0.144	0.095	−6.513	2	1968	1987	0.020	0.017	0.199
	Gansu	−1.423	0.173	0.079	−8.841	7	1971	1974	0.000#	0.000#	0.012
	Qinghai	−0.357	−0.066	−0.067	−3.976	8	1978	1987	0.773	0.805	0.571
	Ningxia	−0.457	0.084	−0.062	−5.470	5	1965	1989	0.143	0.152	0.260
	Xinjiang	−0.349	−0.131	−0.083	−5.352	7	1961	1966	0.178	0.187	0.291

[a]The sign # indicates that the p-value for the province was estimated to be zero due to the fact that the realization of the minimum t_i^m-statistic lay far left from its empirical distribution which was generated by a Monte Carlo simulation with 5,000 replications. Therefore, in order to calculate the Fisher_B statistic, the obtained p-value was assigned a value of 0.0002 (1/5000). This implies that we assume that the minimum t_i^m-statistic took a value within the estimated empirical distribution only once in the 5,000 replications.