Journal of Asia-Pacific Studies (Waseda University) No. 19 (January 2013)

Growth and Convergence among China's Provinces between 1978 and 2003*

Atsuyuki Kato†

This paper examines income and total factor productivity (TFP) convergence across China's provinces for the period 1978–2003, using a fixed-effect panel methodology. Here, we discuss the following questions. (1) Are China's provincial income and TFP converging? (2) Do the heterogeneous provincial TFP levels explain the provincial income gaps? To answers these questions, the Solow growth model is used as the main theoretical framework and the bias corrected Least Squares Dummy Variable approach proposed of Bun and Carree (2005) is employed as our empirical approach. Our results reveal that it is difficult to find evidence of convergence in provincial income levels for the economic reform period of China. The annual convergence rate is only 2.38 percent. The gaps of TFP levels between the coastal and internal regions have increased as well. The relative provincial income levels have more faithfully reflected the relative provincial TFP levels since 1990 than before.

1. Introduction

In this paper, we examine provincial growth and convergence in China. A key question in modern growth theory is whether there is a common steady state to which initially disparate economies converge. In other words, do poor economies catch up with rich ones in the long run? In order to answer this question, much work has been devoted to the empirical analysis of convergence over the last decade and a half. This paper adds to this research by testing for convergence in a panel of 28 provinces of China during the period 1978–2003.

The initial surge of recent growth and convergence empirics started from and focused on cross-country and cross-section analysis, following Abramovitz (1986) and Baumol (1986), partly because of the restriction of data availability. Although those studies have contributed significantly to the issues of growth and convergence, their limitations have also been made clear as the data availability for the regional and panel analysis has improved. With development of the estimators, it has motivated many researchers to focus on panel data and on studies within specific groups of economies, such as OECD countries, low-income countries, transition economies, and so on. Our study is also along these lines. Since China is a low income country with a fifth of the world's population, the regional analysis of its growth and convergence can be considered as a contribution to examining the ideas of growth theory within low-income countries. In addition, it also sheds light on the research of transition economies

* I am grateful to Maurice J. G. Bun (Universiteit van Amsterdam) for allocating the Ox code to estimate his bias corrected LSDV (least square dummy variables).

† Assiatant Professor, Graduate School of Asia-Pacific Studies, Waseda University,
e-mail: akato@aoni.waseda.jp

because China is in transition from socialism to capitalism.

In the empirical study of convergence, following questions are largely examined. First, whether economies converge as a system and if they do at what speed? Second, do the dynamics of total factor productivity (TFP) contribute to the convergence or divergence of income levels? For the first question, we have to define what convergence is, and make it clear how we estimate its speed. The second question requires us to measure TFP levels appropriately.

For the definition of convergence, there are two different concepts, β-convergence and σ-convergence. Both forms of convergence are related, but illustrate different phenomena. β-convergence represents the transition dynamics of economies to their steady states while σ-convergence measures the reduction of the dispersion in income distribution. β-convergence is thought to be the process through which poor economies catch up with rich countries. Therefore, β-convergence is estimated to answer the first question. On the other hand, the analysis of σ-convergence is applied to reveal what forms of convergence occur when economies converges. If σ-convergence exists, β-convergence always exists, but the existence of β-convergence does not always guarantee that σ-convergence exists. The relations of these two convergence concepts are discussed by Young, Higgins, and Levy (2007).

In estimating β-convergence, there are also two different forms, unconditional and conditional convergence. The former is convergence to the common steady state while the latter is to the parallel steady states. Since initial studies largely rejected the unconditional convergence for cross-country analysis, many researchers have paid attention to conditional convergence.[1] As to it, Barro and Sala-i-Martin (1992), (henceforth BS) and Mankiw, Romer, and Weil (1992), (henceforth MRW) formalized the model-based specification. In both works, the regression specification is derived from the neoclassical growth model. The BS specification is based on the Cass–Koopmans' optimal savings version of the neoclassical growth theory while the MRW specification is from the original Solow–Swan model.[2] In addition, Islam (1995) expanded the MRW model to the panel data. Their papers have produced large amount of empirical works. To answer the first question, we also follow Islam's method.

Their contribution has also motivated empirical researchers to answer the second question. Based on their models, recent studies show that the presence of provincially heterogeneous technology or TFP levels is testable in cross-sectional convergence analysis, using appropriate fixed effect estimators. For example, De la Fuente (2002) examines technology diffusion by means of a fixed effect model with independently computed technology levels as a regressor. Islam (2003) and Di Libert, Mura, and Pigliaru (2007), (henceforth LMP) also propose another approach to test whether income convergence is partly due to TFP (including technology) convergence without independent indexes of regional technology levels.

The purpose of the study in this paper is to answer these questions in the case of China's provinces.

1 Rejection of unconditional convergence also motivated to develop endogenous growth models such as Lucas (1988) and Romer (1986).

2 Cass (1965), Koopmans (1965), Solow (1956) and Swan (1956)

Here, the following approaches are applied. First, the descriptive analysis on σ-convergence is undertaken. This provides an overview of the distributional dynamics of Chinese provincial incomes. In addition, it allows us to avoid misunderstandings from the Galton's fallacy which Quah (1993), and Friedman (1992) discussed in the conventional regression approach.[3] Secondly, the standard convergence regression based on MRW and Islam (1995) is examined to estimate the rate of convergence to the steady state. Thirdly we apply the LMP approach to test whether provincial TFP levels form the distribution of provincial income levels.

Compared with the preceding study of China's provincial income convergence, the present study has three advantages. First, we apply the bias corrected least square dummy variables (LSDV) to the estimation of convergence rates in addition to other conventional methods such as ordinary least squares (OLS) and generalized method of moments (GMM). It helps us to discuss issues of convergence rates further. Secondly, the relations between income and TFP convergence are examined in the case of China's provinces. In addition, the present study explicitly discusses the different dynamics of σ-convergence based upon two kinds of GDP deflators. Although such difference gives us interesting interpretations on the convergence mechanism, most of the preceding papers have not paid enough attention to it.

The layout of this paper is as follows. In section 2, we show the model which is examined and discuss the methodologies of estimation. In section 3, we explain the data which we use in this study. Section 4 gives an outline of the provincial and regional income distribution based on descriptive analysis. Section 5 presents the results and a discussion of the convergence estimation. In section 6, the LMP approach is applied to China's empirics. And the last section concludes.

2. Model and Estimation Methodologies

In order to examine the convergence hypothesis, the β-convergence approach based on the Solow growth model is applied. As we mentioned above, β-convergence has the following two forms, unconditional and conditional. Many papers examined unconditional β-convergence in China so far. Therefore, in order to avoid redundancy, the unconditional β-convergence estimation is not repeatedly examined and we focus only on conditional convergence. In this section, we discuss the model which we estimate and advantages and disadvantages of the panel analysis and estimators. First, we provide a brief overview of the Solow Growth model and expand it to the panel data.

Assuming constant returns to scale, we write the neoclassical growth model as follows.

$$Y(t)=K(t)^{\alpha}(A(t)L(t))^{1-\alpha}, \quad 0<\alpha<1 \tag{1}$$

where, Y=the flow of output, K=capital stock, A=labour augmenting technology, L=labour stock, α= capital elasticity, and t denotes the period t.

[3] β-convergence is not a sufficient condition for σ-convergence.

We also assume the constant rate of exogenous technological progress (g) and labour growth (n).

$$A(t)=A(0)e^{gt} \tag{2}$$

$$L(t)=L(0)e^{nt} \tag{3}$$

where $A(0)$ and $L(0)$ are respectively initial states of labour force and technology. Here, we denote labour productivity as $y(t)=Y(t)/L(t)$. For this simple growth model, MRW shows that the logarithmic form of the production function at the steady state is written as follows.

$$\ln y(t)=\frac{\alpha}{1-\alpha}\ln(s)-\frac{\alpha}{1-\alpha}\ln(n+g+\delta)+\ln A(0)+gt \tag{4}$$

where s=investment rate, so, equals to saving rate, and δ is the depreciation rate.

Using equation (4), Islam expands the MRW model to the panel data as follows.

$$\ln y(t_2)-\ln y(t_1)=(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(s)-(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(n+g+\delta)$$
$$-(1-e^{-\lambda t})\ln y(t_1)+(1-e^{-\lambda t})\ln A(0)+g(t_2-e^{-\lambda\tau}t_1) \tag{5}$$

or

$$\ln y(t_2)=(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(s)-(1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(n+g+\delta)$$
$$+e^{-\lambda\tau}\ln y(t_1)+(1-e^{-\lambda\tau})\ln A(0)+g(t_2-e^{-\lambda\tau}t_1) \tag{6}$$

where τ is the years between period t_2 and period t_1. Because of the applicability to different forms of estimators, equation (6) has been much more widely employed in preceding works. We also follow them.

In equation (6), the intercept consists of two terms:$(1-e^{-\lambda\tau})\ln A(0)$ and $g(t_2-e^{-\lambda\tau}t_1)$. As to those terms, in the actual estimation, Islam and Caselli, Esquivel, and Lefort (1996) allow the aggregate production function to vary across countries with respect to $\ln A(0)$ while g is assumed to be identical. In this regression, if the term $\ln A(0)$ is homogeneous across provinces, the differences of the steady states are explained only by saving rates and population growth rates. But if it varies across provinces, it implies that the income gaps are partly from heterogeneity of TFP.

Before estimation, we have to assume the value of sum of the common technological progress rate and the identical rate of depreciation. Since there is no ultimate rule to the assumption, in this study, we follow the assumption that $(g+\delta)$ is equal to 0.05.[4]

Another important issue in the panel analysis of β-convergence is the appropriate time span of the

[4] In some preceding research of China's provincial income convergence such as Raiser (1998) and Weeks and Yao (2003), $(g+\delta)$ is assumed to be 0.07 because China's official depreciation rate is annually 5 percent. Both assumptions give very close results.

observations for transitional dynamics. In order to reduce the effect of fluctuation from the business cycle, the five-year time span is widely used. For example, Weeks and Yao (2003) use five-year average data ($y(t)$ is the series of means of per capita GDP every five years), and Caselli et al. relied on five-year interval data ($y(t)$ is the series of per capita GDP every five years). In this study, we follow the model specification of Caselli et al., and also use five-year intervals. With the linear restriction, we rewrite equation (6) as follows.

$$\ln y_{iT} = \beta \ln y_{i,T-\tau} + \gamma x_{iT} + \eta_T + \mu_i + v_{iT} \tag{7}$$

where $\beta = e^{-\lambda\tau}$, and $x_{iT} = \ln(s_{it}) - \ln(n_{it} + g + \delta)$, $\gamma = (1-\beta)\dfrac{\alpha}{1-\alpha}$, $\mu_i = (1-\beta)\ln A(0)_i$, and $\eta_T = g(t_2 - \beta t_1)$. The saving rate s_{it} and the population growth rate n_{it} are taken as averages between $(T-1)$ and $(T-\tau)$. The term μ_i consists of various unobservable factors (institutions, climate, technology and so on), and is thought to be possibly correlated with other regressors in equation (7). Thus, a fixed effect estimator is thought to be appropriate in this model.

In growth and convergence study, using the panel data model has some advantages. First, we can use a larger size of observations in estimation. Since we have data of only 28 provinces, the single cross section approach has severe limitations in keeping the degree of freedom large enough. The panel analysis can remove this limitation. Secondly, it is possible to deal with issues of both cross section and time series at the same time. Under the cross section analysis, unobservable factors are assumed to be identical through the whole period across provinces. But it seems to be too strong an assumption. In the case of China, the results of estimation are possibly dependent on the choice of the estimated periods since China has not always been politically stable. The panel approach allows us to control this issue to some extent. Time specific shocks are considered as time dummies, and region specific shocks are controlled as individual dummies. Thirdly, the panel approach is most suitable for the model based on dynamics around the steady state since the transition to the steady states is assumed to take place in the short run. Fourthly, we manage to solve the problem of correlation between economy-specific effects and regressors. In the cross-section framework, the estimates are not reliable if omitted unobserved economy-specific effects have correlation with regressors, but we can overcome this problem by using the dynamic panel model with fixed effects.

One disadvantage of the panel approach is that the results are sometimes biased or inefficient because of the estimators themselves and short sample sizes. The properties of dynamic panel data (DPD) estimators are typically investigated for microeconomic datasets with a large number of individuals (N) but a small time dimension (T). However, differences of the characteristics between microeconomic and macroeconomic datasets possibly influence the robustness of estimation techniques which are used. Therefore, the simulations by Monte Carlo experiments have been undertaken, in order to compare the robustness of some different estimators. Judson and Owen (1999), and Islam (2000) examine the properties of DPD estimation for macro dataset. Gaduh (2002) also examines it for the modified parameter values to capture the typical characteristics of the Penn World Tables (PWT).

They show that there is no hard and fast rule as to which estimation technique is best.[5]

In this paper, we also conduct a Monte Carlo simulation to make clear the properties of those estimators in similar size of data to our study. The simulation design for the model with an additional explanatory variable is basically the same as Bun and Carree (2005) but has different values of observations and parameters. The details and results are presented in Appendix 2. There, Kiviet, Bun and GMM are respectively the bias corrected LSDV by Kiviet, the bias corrected LSDV by Bun and Carree and the GMM by Arellano and Bond. For the cases of (N, T)=(40, 6) and (N, T)=(30, 8), which are similar to our sample size, Bun and Carree-corrected LSDV provides less biased estimates of the coefficient of β. The root mean square errors are also relatively smaller where β is close to unity. As to GMM, the bias for estimates of β increases as the number of individuals decreases. It indicates that the first difference GMM suffers larger downward bias in the small sample.

In addition to Monte Carlo simulation above, the results of Monte Carlo simulations in the previous paper show that LSDV perform better than the Anderson and Hsiao IV estimator and the Arellano and Bond first difference GMM estimator for the small sample dynamic panel. It indicates that the bias corrected LSDV deserves to be applied to the small sample dynamic panel. On the other hand, the GMM methods should also be examined because the problem of endogeneity is not negligible in growth models. Thus, we examine the following estimators: Kiviet-corrected LSDV (1995), Bun and Carree-corrected LSDV (2005), the first difference GMM, and the system GMM. In addition, the estimates of the OLS and the within group estimator are used as upper and lower bounds respectively.

3. Data Issues

One of the big issues in empirical studies of developing countries is the severe limitation on the availability of data. In some countries, poor availability of statistics gives serious disadvantages to empirical studies. In addition, the comparability and reliability of data are often in questions in transition economies. Since China is a transition economies and our empirical study largely relies on officially published statistics, it is important to make clear the sources and the properties of them before discussion.

The official statistics of China are accumulated and published by the National Statistical Bureau of China (NSB). During the pre-reform period, the Chinese authority had followed the Material Production System (MPS) as the former Soviet Union had done.[6] But under their economic reforms, China has also reformed its national statistic system from the MPS to the System of National Accounts (SNA) and re-constructed data even for the pre-reform period, supported by the World Bank. Although those data still have some problems,[7] they are consistent and most reliable at the Macro-economy level, in

5 As DPD estimators, OLS, LSDV.

6 The MPS was widely used in the communist economies and doesn't include the non-material service sectors because it is based on the Marxian economic theory.

7 The data problems are discussed in Lin (2000). But those problems are not thought to severely harm the validity of analysis based upon such official statistics.

particular, for the reform period. This newly constructed data have motivated a number of research using familiar econometric methodologies since empirical work relied on them is thought to be comparable to empirical work for other countries and regions. Our paper also relies on such data.

Our data consists of 28 China's provinces including the three municipalities (Beijing, Tianjing, and Shanghai) as provincial level jurisdictions between 1978 and 2002.[8] Due to absence of data, Hainan and Tibet are excluded.[9] Chongqing is included in Sichuan province because Chongqing was separated from Sichuan in 1996. GDP and investment data are obtained from various issues of the China's Statistical Yearbook, and Huesh and Li (1999). Implied deflators of GDP and investment are also from the same sources. Population and retail price index are from "Comprehensive Statistical Data and Materials on 50 Years of New China" compiled by the NSB, and also various issues of the China's Statistical Yearbook.

The NSB provides only nominal data. Therefore, we have to construct the series of real values of statistics by ourselves. For GDP, we have two different deflators, the implicit deflator, and the retail price index. The implied deflators are constructed by GDP at current yuan and volume indices at "comparable" (quasi-constant) prices. This deflator is based on production, and suitable for an analysis of productivity. On the other hand, the retail price index is assumed to be a substitute of the consumer price index (CPI),[10] and suitable for an analysis of average income. As to investment, the official investment deflators are available only from 1991. We, therefore, have to rely on the implicit deflators. As to Guangdong province, due to the absence of data, the index of national average is used as a substitute between 1996 and 2000.

4. Descriptive analysis

Descriptive analysis gives us an overview and some stylized facts about China's economic growth and convergence. Table 1 shows the provincial (also regional) income levels and growth rates.[11] From the table we learn the following. First, the ranks of provincial income levels are not stable before and after 1978. It implies that some decisive factors or the structure to form distribution of provincial average income is not always the same before and after 1978. Secondly, in all provinces, their growth rates are higher in the reform period than in the pre-reform period. However, the acceleration of growth rates is highly heterogeneous across provinces. Twelve provinces experienced higher appreciation of their growth rates than the average, and they consist of four eastern, four central and four western provinces. On the other hand, five eastern including three municipalities, five central and six western provinces suffered lower appreciation of their growth rates than the average. It suggests that there is different dynamics even within the same regional group.

---

8 The main land China consists of 31 provinces excluding Hong Kong and Macao.
9 The shares of their population and GDP are less than 1 percent in national total.
10 The CPI doesn't cover the whole period for some provinces even in the reform period.
11 Province code and location are in Appendix 1.

Table 1 Provincial Per Capita Incomes and Growth Rates

Province Code	Income 1953	R	Income 1978	R	Income 2002	R	Growth 1953-1977	Growth 1978-2002	Growth Difference
BJ	1558.16	1	3323.06	2	19344.12	3	2.80%	7.34%	4.54%
TJ	1009.12	4	2990.87	3	19651.03	2	3.80%	7.84%	4.04%
HE	374.5	20	1074.22	13	8589.57	11	3.88%	8.66%	4.78%
SX	441.04	17	1079.35	12	6082.83	19	3.10%	7.20%	4.10%
NM	581.97	8	941.17	20	6892.4	15	1.74%	8.30%	6.56%
LN	647.31	6	1976.79	5	12048.1	8	4.29%	7.53%	3.24%
JL	539.12	9	1118.06	10	8072.64	12	2.60%	8.24%	5.64%
HL	1051.64	3	1998.16	4	9752.63	10	2.31%	6.61%	4.30%
SH	1429.89	2	5133.95	1	31707.31	1	4.76%	7.59%	2.83%
JS	504.81	10	1148	7	14590.04	5	2.56%	10.59%	8.03%
ZJ	450.94	14	1040.31	14	15871.44	4	2.71%	11.35%	8.64%
AH	480.49	11	757.27	24	6122.74	18	1.95%	8.71%	6.76%
FJ	463	13	985.53	17	13433.46	7	2.54%	10.88%	8.34%
JX	465.48	12	734.49	26	5694.76	22	1.47%	8.53%	7.06%
SD	327.9	25	1030.18	16	11382.31	9	4.40%	10.01%	5.61%
HA	409.94	18	734.96	25	6026.44	20	2.05%	8.77%	6.72%
HB	450.36	15	966.8	18	7956.88	13	2.71%	8.78%	6.07%
HN	442.32	16	1031.07	15	6300.72	17	2.94%	7.54%	4.60%
GD	582.34	7	1109.44	11	13723.21	6	2.73%	10.48%	7.75%
GX	346.98	23	958.73	19	5628.45	23	3.95%	7.37%	3.42%
SC	352.37	22	802.98	21	5553.52	24	2.35%	8.06%	5.71%
GZ	321.93	26	538.1	28	3004.5	28	1.33%	7.17%	5.84%
YN	369.23	21	788.21	22	4367	26	2.43%	7.13%	4.70%
SN	298.36	27	775.29	23	4937.68	25	3.59%	7.71%	4.12%
GS	334.01	24	723.04	27	4040.31	27	2.75%	7.17%	4.42%
QH	406.46	19	1484.58	6	6368.02	16	4.91%	6.07%	1.16%
NX	290.19	28	1119.84	9	5933.64	21	5.38%	6.95%	1.57%
XJ	676.07	5	1125.63	8	7432.75	14	1.83%	7.86%	6.03%
EAST	699.54	11	1888.28	10	15088.1	7	3.49%	9.06%	5.57%
CENTRE	515.25	14	1048.12	16	6883.57	17	2.62%	7.90%	5.28%
WEST	394.06	21	891.12	19	5100.54	23	2.74%	7.31%	4.57%

Despite such complexity in both of the average income levels and growth rates, the ranks of the three regions, east, centre, and west have remained unchanged. Rather, the average of the east region is reinforced, in particular in the reform period. This indicates that there are some structural causes to

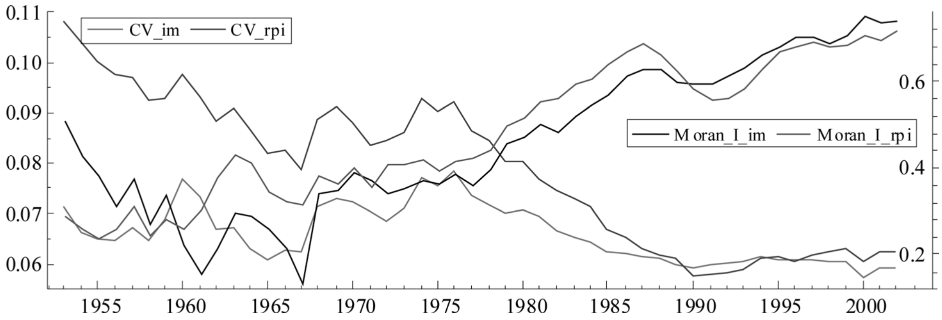

Figure 1 Coefficient Variance and Moran's I statistics

Note: As to the RPI., the following data are missing.

Henan,	[1953–55]
Ningxia	[1953–56]
Anhui Jiangxi	[1953–69]
Gansu Ningxia	[1953–70]
Guangxi Sichuan	[1953–76]

Coefficient Variance and Moran's I statistics: author's calculation

form the regional gaps in the reform period. Figure 1 draws trends of σ-convergence (CV) and Moran's I statistics.[12] Since some previous studies are based on per capita GDP deflated by the retail price index (RPI), the figure shows both CVs on the implicit deflators and the RPI. The figure clearly shows that there is σ-convergence in the 1980s. Some preceding studies such as Jian, Sachs and Warner (1996) or Raiser (1998) also used the RPI as deflators and find convergence in the 1980s. But in the case of the implicit deflators, the scale of σ-convergence is almost a half of that in the case of the RPI.

As is stated above, the GDP per capita deflated by the implicit deflators from the comparable prices is based on production while that deflated by the RPI is thought to be suitable for the average income. Therefore, the disparity of the CVs based on those deflators indicates the differences of distribution between production (or productivity) and income. During the pre-reform period, the variance based upon the RPI is consistently larger than that based on the implicit deflator.[13] This means that the gaps of income are exaggerated more than the gaps of their production. Therefore, it is clear that the egalitarianism strategy in the pre-reform period couldn't reduce the provincial gaps of incomes which stemmed from the gaps of production, and rather increased the inequality between the municipalities and other provinces.[14]

---

[12] Moran's I statistics indicate spatial correlation, and calculated as follows.

$$I = \frac{N\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} z_i z_j}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} \sum_{i=1}^{N} z_i^2}$$

where N=the number of values to be taken into account, yi/j=value at location i and j, w_{ij} is the weight at distance d, and z's are deviations (i.e. $z_i = yi—y$ mean for variable y).

[13] The absence of data in some provinces doesn't seem to severely affect the result because the CV from 21 provinces (smallest availability during the pre-reform period) is considerably close to that from all 28 provinces for the reform period.

[14] The gaps of CVs excluding the municipalities are much smaller than those including them but not carried in this study.

The reason is thought to be as follows. As Minami (1994) said, in the pre-reform period, the Chinese government kept the prices of agricultural products very low and the prices of industrial products high, to increase the surplus and investment in industrial production. Therefore, in terms of the RPI, the income of rural farmers was lowered while the income of urban industrial workers was raised. Since the municipalities have considerably different structures of industry and employment, and focus on urban industrial production,[15] the gaps in the case including those municipalities are exaggerated.

In addition to this reason, the self-reliance policy under Mao's regime possibly worked to increase gaps. The gaps during the period of the Great Leap Forward (1958–60) and the Cultural Revolution (1965–76) increased. In these periods, division of labour, comparative advantage and economic efficiency were ignored because of Maoists' ideology and the militaristic strategy as we state in the previous chapter. Therefore, it was unlikely to equalize productivity and income across different provinces through spillover effects. Rather, such policies are supposed to increase the gaps by the reason mentioned above.

In the same way, σ-convergence which we find in CVs of the 1980s is explained by adjustment of politically controlled prices to the market prices to some extent. The agricultural reform (abolition of people's communes and adoption of the household responsibility system) and marketisation are thought to reduce such price distortion to some extent until 1990.[16]

In Figure1, Moran's I statistics indicate the degree of spatial correlation among provinces. The resultant values of computing Moran's I statistics fall in the range from approximately −1 to 1. The positive value of it represents positive spatial autocorrelation, while the converse is true for the negative value. If the result is zero, it represents no spatial autocorrelation. It is obvious that regional correlation fluctuated in the pre-reform period but it has been gradually reinforced in the reform period.

In the pre-reform period, the low values of Moran's I statistics based on the implicit deflators during both the Great Leap Forward and the Cultural Revolution are consistent with the facts that the self-reliance policy cut the inter-regional interactions. But the controlled price seemed to absorb the effects of such political shocks on average income to some extent. As to the reform period, since the values of the statistics have increased while the CVs reached a stable level of a-spatial inequality, it is thought that polarization at constant levels of inequality has developed in the 1990s.[17] However, such a high degree of spatial correlation is not from increasing inter-provincial interactions because inter-provincial migration is still strictly limited and inter-provincial trade is, rather, decreasing according to Young (2000). Therefore, in order to explain such polarisation, it is required to examine what factors crucially work using regression analysis of conditional convergence.

---

[15] More than half of the population are classified as non-agricultural population in those municipalities while almost two thirds of population are classified as agricultural population on the national average.

[16] This doesn't mean that the implicit deflators directly reflect the market prices. China still relies on controlled prices in many areas, and the implicit deflators are dependent on strong assumption on comparable prices: in base year, comparable prices of standard products are assumed to be identical across all regions.

[17] Arbia (2001).

Table 2 Panel Data Test for Conditional Convergence of the Reform Period 1978–2003

Dependent Variable: ln $y(t_2)$
Upper Bound (OLS): 1.003
Lower Bound (Within): 0.733

	Bias Corrected LSDV				GMM			
	(1) Kiviet BC		(2) Bun BC		(3) DIFF-GMM		(4) SYS-GMM	
$\ln y(t_1)$	0.955	***	0.888	***	0.688	***	0.983	***
	(0.059)		(0.080)		(0.087)		(0.035)	
$\ln(I/GDP) - \ln(n+g+\delta)$	0.032		0.040		0.048		-0.003	
	(0.046)		(0.045)		(0.064)		(0.063)	
Constant	N/A		N/A		0.225	***	0.459	
					(0.030)		(0.250)	
Implied α	0.496		0.480		0.420		N.A	
Implied λ	0.94%		2.38%		7.48%		0.34%	
M2					0.120		0.116	
Sargan Test					0.675		0.738	

Note: values in parenthesis are standard errors
***, **, * mean 1%, 5%, and 10% statistical significance, respectively.

α is calculated as follows: $\alpha = \frac{\gamma}{(1-\beta+\gamma)}$

***, **, * mean 1%, 5%, and 10% statistical significance, respectively.

α is calculated as follows: $\alpha = \frac{\gamma}{(1-\beta+\gamma)}$

5. Results of Convergence Estimation

In this section, we discuss the results of convergence estimation. Table 2 shows the result of estimations of per capita GDP convergence in various methods, respectively.[18] The OLS and LSDV estimates of the coefficients on the lagged dependent variable (β in equation (7)) are used as the upper and lower bounds of β because OLS and LSDV estimate the upward and the downward biased βs in the fixed effect dynamic panel with short time series, respectively. Since the estimate of β in the first difference GMM is lower than the lower bound, it is thought that the first difference GMM suffers severe downward bias as we discussed in the previous section. On the other hand, estimates in the system GMM and the bias corrected LSDV fall between the bounds. Therefore, we pay attention to interpreting the

[18] All results are estimated using the DPD software for Ox. See Doornik, Arellano, and Bond (2002).

results of such rather than the result of the first difference GMM. The Sargan test statistics indicate that the validity of the moment conditions is not rejected.

The rates of convergence in the two bias corrected LSDV and the system GMM are annually 0.94, 2.38, and 0.34%. It means that it takes 73.7, 29.1, and 203.9 years to reduce the gap to the half level, respectively. Between the two bias corrected methods above, Bun and Carree-corrected estimator shows better performance in estimating the coefficient on the lagged dependent variable (less biased β), according to our Monte Carlo simulation. Thus, we also rely on it and conclude that the annual convergence rate to their steady states is 2.38 percent.[19] On the other hand, the result of the system GMM shows that there is almost no convergence as a system.

One problem to which we should pay attention is that the coefficients on the lagged dependent variable in columns 1 and 4 imply the possible presence of unit roots because both of them are close to unity (0.954 and 0.983). Between them, the system GMM is thought to avoid problems related to unit roots such as spurious correlation because it uses the first differences as well as the levels. To examine robustness of estimation, we impose a restriction of a unit root in equation (7) and estimate it. That is, imposing the restriction that $\beta=1$, the lagged dependent variable ($\ln_{t,T-\tau}$) is transposed to the left side of equation (7).

$$\Delta \ln y_{iT}=\gamma x_{iT}+\eta_T+\mu_i+v_{it} \tag{7'}$$

where $\Delta \ln y_{iT}=\ln y_{iT}-1\, y_{i,T-\tau}$. The result of estimation is similar to that of equation (7).[20] On the other hand, for bias-corrected LSDV, it possibly depends on a technique chosen because a less biased estimate by Bun and Carree's approach doesn't always imply the presence of unit roots. Therefore, it is not thought that this issue gives serious problems to the result of our estimation.

The restricted explanatory variable is not statistically significant, either. In the same model, if we loosen the linear restriction on explanatory variables, both the population growth and the saving rate are statistically insignificant as well.[21] In addition, the system GMM estimation without the restriction gives a negative estimate of coefficient on the saving rate while it is expected to be positive from the model. It implies that it is difficult to detect the crucial role of capital deepening to decide the performance of provincial growth and convergence in the reform period using the Solow–Swan model. The following reasons are considered to explain this finding.

First, the technological catching up might be important to explain growth and convergence dynamics, rather than capital deepening, in China for the reform period. The assumption of a strictly exogenous common technological progress rate and an identical depreciation rate, that is, $(g+\delta)$ is equal to 0.05,

[19] If we examine the same model using the four-year interval for the similar period, 1978–2002, the estimated convergence rate is annually 2.36%.

[20] The estimation results are following:
$\gamma=-0.068$ (0.096), $\mu=0.425$ (0.139). The Sargan Statistics (p-value) is 0.591.

[21] The estimation results (the system GMM) are following:
$\beta=1.005$ (0.041), $\gamma_n=-0.255$ (0.290), $\gamma_s=0.006$ (0.148), $\mu=-0.474$ (0.762), where γ_n and γ_s are coefficients of $\ln(n_{it}+g+t)$ and $\ln(S_{it})$. The Sargan statistics (p-value) is 0.960, and the restriction (p-value) is 0.224.

is necessarily in order to discuss the average convergence rate in a system of economies under the framework of MRW, but not always reasonable. Since the accessibility to the advanced foreign technology through export-oriented investment is different between the coastal and inland provinces, it is possible that the technological progress rates are also heterogeneous even though we still hold the assumption of an identical depreciation rate.

Secondly, the quality of the statistics is controversial. The statistics of gross capital formation in China Statistical Yearbook doesn't mean the completed capital formation in the period. The statistics include all capital formation which is just planned. Thus, "investment" is not always activated for production in the reported periods. In particular, investment by state-owned companies often has more than one-year plans because of the five-year development plans by the central government, and is often inefficient due to the soft budget problems. This problem is beyond our control, but we consider that it is not a crucial issue because we use the five-year average.

6. Growth and Technology Convergence

In this section, we discuss TFP catching up as a major source of growth and convergence dynamics of China's provinces. Many theoretical and empirical models are employed to incorporate technological progress with growth and convergence dynamics so far. Among them, we examine one of the simplest methods proposed by Di Liberto et al. and Islam. The reason is following. Many of the empirical models for discussing technological catching up are dependent on individual indexes of regional technology levels, but they are not always available in developing countries. Compared with them, the method which is examined in this paper uses only the individual intercepts to compute an estimate of the total factor productivity (TFP) levels from a result of LSDV. Because of the severe downward bias in the LSDV, we use the Bun's bias correction method as well as the estimation in the last section. Since this method is applied to the conventional Islam's growth model for panel data which we applied in the last section, it is suitable to compare their results.

This method gives no information to the internal mechanism of technological (productivity) convergence, but that is not in the scope of this section. Here, we focus only on discussing whether we can detect technological convergence, and whether the dynamics of technological convergence is consistent with the facts from the descriptive analysis, and reinforce the discussion in the last section if we find it.

The framework of methodology and ideas are following. First, we divide the regression period into two sub-periods, initial and subsequent. Secondly, we estimate equation (7) in both periods by the bias corrected LSDV. Thirdly, we compute estimates of TFP levels from the individual intercepts. Fourthly, we compare the relative TFP levels between the initial and the subsequent periods. If there is no technological catching up, the distributions of the relative TFP levels in initial and subsequent periods are identical. Otherwise, there is convergence or divergence in technology levels.

In the last section, we assume that μ_i is constant in order to discuss effects of capital deepening on convergence. But, in this section, we assume μ_i is a time-invariant component that varies across economies. The term μ_i includes various unobservable factors, such as institutions, climate, or technology. Since many of such unobservable factors are assumed to affect a desirable technology, we consider the TFP level as the technology level. In addition, technology levels seems to be correlated with the original regressors, in particular, the lagged GDP per capita in the model, the fixed effect estimator is appropriate. From $\mu_i=(1-\beta)\ln A(0)$, it is easy to compute a proxy of TFP if we have the estimated individual intercepts. Since we can't directly obtain the estimate of μ_i in the estimation of the bias corrected LSDV, the following procedure has been used to obtain the estimates indirectly.

$$(\hat{\mu}_i+\hat{v}_{iT})=\tilde{y}_{iT}-\hat{\beta}\tilde{y}_{iT-1}-\hat{\gamma}\tilde{x}_{iT} \tag{8}$$

$$\hat{\hat{\mu}}_i=\frac{1}{T}\sum(\hat{\mu}_i+\hat{v}_{iT}) \tag{9}$$

where $\tilde{y}_{iT}, \tilde{y}_{iT-1}, \tilde{x}_{iT}$ are respectively $y_{iT}-\bar{y}_T$, $y_{iT-1}-\bar{y}_{T-1}$, $x_{iT}-\bar{x}_T$. The bar means average and the hat is the sign that it is obtained from estimation.

The next point is how we detect technological convergence in this empirical framework. In order to answer this question, we define the over time evolution of $\ln A_i$ and remind the assumption in Islam's approach. Between period 0 and T, the level of technology in economy i at time T is equal to $\ln A_{iT}=\ln A_{i0}+\gamma_i T$ where $\gamma_i T=gT+\rho_i T$. Here, g is defined as the long run technological progress rate, and assumed to be constant across economies. If there is technological catching up, $\rho_i T$ is different from zero, and a positive function of the technological gap at initial period. On the other hand, under the assumption of original Islam's model, $\rho_i T=0$. In this case, heterogeneity of technology level is based on the stationary gaps of technology, and all economies have a common rate of technology growth. To compare these two cases, we investigate the distribution of technological gaps, which is defined as A_i/A^* where A^* is the technology level of the leader province, and A_i is the same index for a follower. If the Islam's assumption of technological progress that the difference of technology levels is stationary is accepted, then, A_i/A^* is constant. To a contrary, if there is technological convergence, then, A_i/A^* is increasing over time.

We examine this test based on the similar data to the estimation in the last section.[22] Di Liberto et al. and Islam add education attainment as a proxy of human capital to the basic model, but we don't. The reason is following. Even though human capital is a theoretically important factor of economic growth, it is not clear whether education attainment is an appropriate proxy for it. Human capital consists of many factors, and education is just part of them. In addition, unlike Italian case in Di Liberto et al., the relationship between growth and education is not always clear in China.[23]

One important point in this analysis is whether we justify the differences of the relative TFP levels between the initial and the sequent periods are statistically significant. In the framework of this analy-

[22] In this section, we estimate equation (7) using four-year interval data between 1978–2002.

[23] Lin(2001) argues this issue.

sis, it means whether we detect structural changes between both periods. In order to examine it, we also conduct the Wald test. The result of the test is that the Wald statistics is 5.1735 and its *p*-value is 0.0753. This result indicates that we don't detect the structural change at the five percent significant level, but find it at the ten percent level. It means that the hypothesis that the technological gaps are stationary is weakly rejected. Therefore the analysis based on the estimated individual effects still seems to be meaningful.

Table 3 shows the relative income and the estimated relative TFP levels. Since Shanghai has kept the top rank in the income level since 1978, we use Shanghai as the leader economy in the above procedure. From the table, we detect some interesting features. Increasing or decreasing of their relative TFP levels is geographically characterized. Among all provinces except for the leader, Shanghai, only eight provinces improved their relative TFP levels. It means that only a third of provinces could ride on the catch-up trend of TFP levels. The number is smaller than thirteen in the relative income catching up. Seven to those eight provinces belong to the eastern region, and one (Jiangxi) belongs to the central, but adjoins the southeast coastal provinces. Therefore it is obvious that only the southeast coastal provinces enjoyed improving their relative TFP levels.

On the other hand, even in the eastern region, Beijing, Tianjin and Liaoning deteriorated their relative TFP levels. In addition, two of the northeast provinces, Jilin and Heilongjiang also felt down their relative positions of TFP levels. It implies that the northeast region failed to capture the convergence trend of productivity. All the western provinces decreased their relative TFP levels, and in most of them, the deterioration of the relative TFP levels is severer than that of the relative income.

The above findings are roughly consistent with the finding from Table 1 that the position of the eastern region is reinforced in the reform period because of rapid growth of the southeast coastal provinces. In addition, the fact that the trends of adjoining provinces are generally similar is consistent with the polarization which is indicated by the Moran's I statistics.

Figure 2 indicates the same results. The horizontal axis measures the relative TFP levels in the initial period, and the vertical does in the subsequent period. The upward slope is the 45-degree line. If the Islam's original assumption (technology gaps are stationary) is accepted, the scatters have to be on or near the line. But, the scatters look dispersed. The coastal provinces (Jiangsu, Zhejiang, Fujian, Shandong, and Guangdong) form one group above the line, while the western provinces are scattered around the low-left corner. In addition, the coastal provinces and the northwestern provinces (Qinghai, Ningxia, Xingjiang) seem to be respectively on the parallel lines over and beneath the 45-degree line.

The next discussion point is the relationship between the distribution of provincial per capita GDP growth (or income levels) and the TFP (technological) progress. The correlation coefficient between the improvement of relative provincial incomes during the reform period and the rate of improvement in the relative TFP levels is 0.8091, and the Spearman's rank correlation coefficient is 0.8431. It means that the difference of technological progress across provinces is highly correlated with the distribution

Table 3 Relative Income and Relative TFP level (Shanghai = 1.0000)

	Income		**TFP**				**TFP**	**Rank**	
Province	**Period 1**	**Period 2**	**Period 1**	**R**	**Period 2**	**R**	**Improve**	**Change**	**Region**
BJ	0.6944	0.6480	0.5083	2	0.5812	2	0.0729	0	East
TJ	0.6018	0.5824	0.3650	3	0.5212	3	0.1562	0	East
HE	0.2139	0.2602	0.0484	18	0.2181	11	0.1698	7	East
SX	0.2255	0.2001	0.0546	16	0.1569	17	0.1022	-1	Centre
NM	0.2160	0.2150	0.0577	15	0.1682	15	0.1105	0	Centre
LN	0.4012	0.3877	0.1654	4	0.3301	8	0.1647	-4	East
JL	0.2471	0.2532	0.0649	12	0.2074	12	0.1425	0	Centre
HL	0.3764	0.3177	0.1418	5	0.2588	10	0.1171	-5	Centre
SH	1.0000	1.0000	1.0000	1	1.0000	1	0.0000	0	East
JS	0.2903	0.4238	0.1083	7	0.3957	6	0.2875	1	East
ZJ	0.2974	0.4615	0.1256	6	0.4400	4	0.3144	2	East
AH	0.1741	0.1912	0.0356	20	0.1561	18	0.1205	2	Centre
FJ	0.2485	0.3939	0.0811	10	0.3893	7	0.3083	3	East
JX	0.1602	0.1768	0.0293	25	0.1430	21	0.1137	4	Centre
SD	0.2421	0.3329	0.0693	11	0.3039	9	0.2346	2	East
HA	0.1739	0.1897	0.0339	23	0.1501	19	0.1163	4	Centre
HB	0.2251	0.2436	0.0598	14	0.2024	13	0.1426	1	Centre
HN	0.2063	0.2005	0.0448	19	0.1616	16	0.1168	3	Centre
GD	0.2802	0.4259	0.1075	8	0.4163	5	0.3088	3	East
GX	0.1647	0.1827	0.0273	26	0.1467	20	0.1194	6	East
SC	0.1775	0.1784	0.0347	22	0.1367	24	0.1020	-2	West
GZ	0.1196	0.0991	0.0159	28	0.0630	28	0.0471	0	West
YN	0.1704	0.1665	0.0349	21	0.1167	25	0.0818	-4	West
SN	0.1716	0.1570	0.0335	24	0.1079	26	0.0745	-2	West
GS	0.1370	0.1405	0.0195	27	0.1025	27	0.0830	0	West
QH	0.2606	0.1984	0.0636	13	0.1412	22	0.0776	-9	West
NX	0.2263	0.1945	0.0538	17	0.1373	23	0.0835	-6	Centre
XJ	0.2668	0.2663	0.0896	9	0.1979	14	0.1084	-5	West
East	0.4031	0.4635	0.2369	8.7	0.4311	6.9	0.1942	1.8	
Centre	0.2225	0.2179	0.0576	16.5	0.1737	16.3	0.1160	0.2	
West	0.1862	0.1723	0.0417	20.6	0.1237	23.7	0.0820	-3.1	

of income levels. In addition to the facts that the capital deepening is not statistically significant in the growth model estimation, and provincial TFP levels are estimated with statistical significance, such high correlation seems to support the view that technological progress has played a crucial role to form growth and convergence performance during the reform period in China.

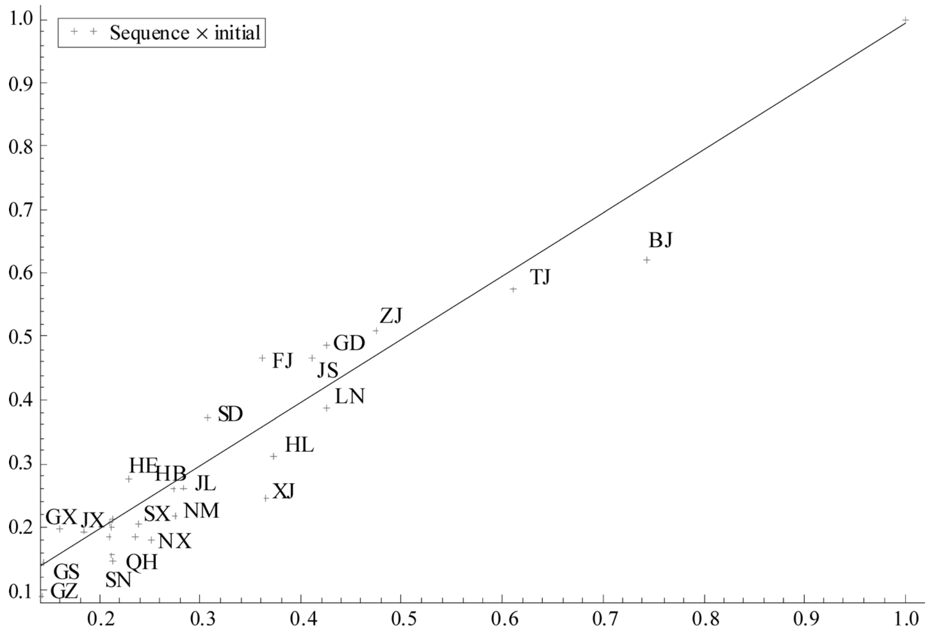

Figure 2 Relative Productivity Dynamics

Note: The upward line is 45-degree line.
Horizontal = relative TFP levels in the Sub-Period 1, Vertical = relative TFP levels in the Sub-Period 2
Note: The upward line is 45-degree line.
Horizontal = relative TFP levels in the Sub-Period 1, Vertical = relative TFP levels in the Sub-Period 2

Figure 3 and 4 also show the relationship between the average income and TFP level for each province in the initial and the subsequent periods, respectively. The initial period basically covers the 1980s and the subsequence does the 1990s. In both figures, the upward lines are also 45-degree lines. From the figures, the change of distribution is obvious. In the 1980s, the relative income levels are generally higher than the relative TFP levels. It implies that there are some important factors to form the distribution of income levels, other than the TFP levels. On the other hand, the levels of income almost follow the levels of the TFP in the 1990s. It indicates that the TFP became a decisive factor to form the distribution as the reform progressed. This finding is also consistent the finding from Figure 1 that exaggeration of income gaps over gaps of production gradually decreased during the reform period, and reinforces the above view on the role of technological progress in the growth and convergence process.

The above analysis doesn't give any direct answer to the question how much of convergence that we observe is due to contribution of technological progress or capital deepening. However, it shows the relationship between the TFP (technology) levels, and growth and levels of per capita GDP in the framework of the standard Solow growth model approach. Adding to the result that investment rate is not statistically significant in the estimation, it seems to support the view that the technological progress is the principal source to form dynamics of provincial growth and convergence.

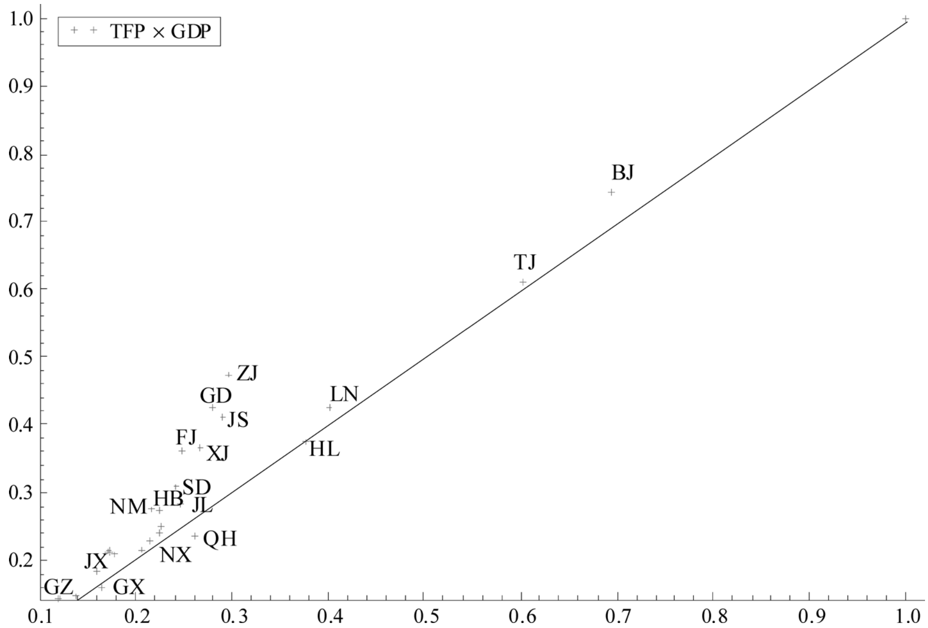

Figure 3 Relative Income and Relative TFP Level in the 1980s

Note: The line is 45-degree line.
Horizontal = Relative Income, Vertical = Relative TFP Level, Correlation Coefficient = 0.9691
The squared sum of deviation from the 45-degree line is 0.1136.

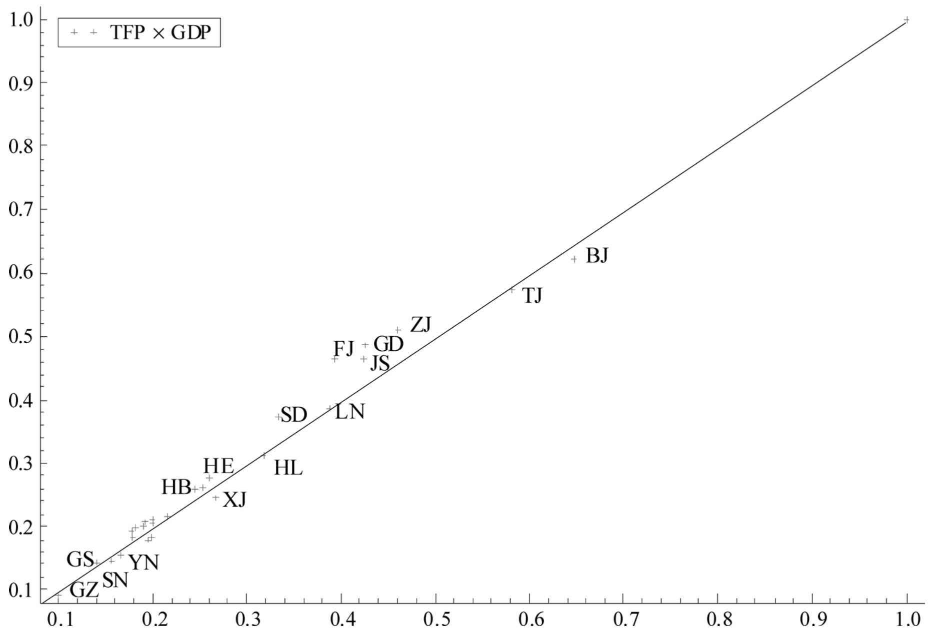

Figure 4 Relative Income and Relative TFP Level in the 1990s

Note: The line is 45-degree line.
Horizontal = Relative Income, Vertical = Relative TFP Level, Correlation Coefficient = 0.9926
The squared sum of deviation from the 45-degree line is 0.0018.

7. Cocluding remarks

In this paper, we discuss the following. Are China's provincial income and TFP converging? And do the heterogeneous provincial TFP levels form the provincial income gaps? From the above discussion, we find the following answers to those questions. First, it is difficult to say that the income and TFP levels of China's provinces are converging as a system. Rather, geographical heterogeneity has been reinforced. The estimated speed of income convergence to their own steady state is annually 2.38 percent. It indicates that it costs long time to reduce the gaps to their steady states. The coastal provinces (Jiangsu, Zhejiang, Fujian, Shandong, and Guangdong) have caught up with Shanghai in terms of the relative TFP levels while most of internal provinces have remained behind. In addition, two municipalities (Beijing and Tianjin) have deteriorated their TFP levels as well. Secondly, the principal source to form income distribution during the reform period is productivity progress, rather than capital deepening. The relative levels of GDP per capita in the 1990s are largely reflected by the relative TFP levels. Our results strongly suggest that it is important to examine what decisive factors are for TFP progress in growth empirics of China.

Appendix 1: Province Code and Location

Province	Code	Location
Beijing	BJ	East
Tianjin	TJ	East
Hebei	HE	East
Shanxi	SX	Centre
Mongolia	NM	Centre
Liaoning	LN	East
Jilin	JL	Centre
Heilongjiang	HL	Centre
Shanghai	SH	East
Jiangsu	JS	East
Zhejiang	ZJ	East
Anhui	AH	Centre
Fujian	FJ	East
Jiangxi	JX	Centre
Shandong	SD	East

Province	Code	Location
Henan	HA	Centre
Hubei	HB	Centre
Hunan	HN	Centre
Guangdong	GD	East
Guangxi	GX	East
Hainan	HI	East
Sichuan	SC	West
Guizhou	GZ	West
Yunnan	YN	West
Tibet	XZ	West
Shaanxi	SN	West
Gansu	GS	West
Qinghai	QH	West
Ningxia	NX	Centre
Xingjiang	XJ	West

Note: The codes follow the Guobiao 2260–1999 alphabetic province codes, issued by the Standardization Administration of China (SAC).

Appendix 2: The Monte Carlo Experiments

[A2–1] Model Settings

In order to examine the properties of various estimators, The Monte Carlo experiments are conducted. The examined model is as follows.

$$y_{it} = \beta y_{i,t-1} + \gamma x_{it} + \mu_i + \varepsilon_{it}, \quad i=1, \ldots, N, \quad t=1, \ldots, T \tag{A1}$$

where μ_i and ε_{it} are independently, identically and normally distributed random variables, $\mu_i \sim IIN(0, \sigma_\mu^2)$ and $\varepsilon_{it} \sim IIN(0, \sigma_\varepsilon^2)$. The initial observation is assumed as follows.

$$y_{i0} = \frac{1}{1-\beta}\mu_i + \frac{\varepsilon_{i0}}{\sqrt{1-\beta^2}}, \quad i=1, \ldots, N \tag{A2}$$

The additional explanatory variable x_{it} is constructed as follows.

$$x_{it} = \rho x_{i,t-1} + \xi_{it}, \quad i=1, \ldots, N, \quad t=1, \ldots, T \tag{A3}$$

where ξ_{it} is also an independently, identically, and normally distributed random variable, $\xi_{it} \sim IIN(0, \sigma_\xi^2)$. We assume the panel data set has 240 observations and conduct experiments for the following combinations of N and T for which $NT=240$; (80, 3), (60, 4), (40, 6), (30, 8), (20, 12). We also assume that $\beta=0.6$, 0.8, 0.9, or 0.95, $\gamma=1$ and $\rho=0.8$. In addition, $\sigma_\eta=\sigma_\varepsilon=\sigma_\zeta=1$ is chosen as well. For each experiment, we perform 10,000 Monte Carlo replications.

[A2-2] Results

(N, T)	(80, 3)	(60, 4)	(40, 6)	(30, 8)	(20, 12)
β=0.95			**Bias β**		
LSDV	-0.180	-0.119	-0.066	-0.042	-0.023
Kiviet	0.003	-0.001	-0.002	-0.002	-0.002
Bun	0.003	0.001	0.000	0.000	0.000
GMM2	-0.013	-0.017	-0.021	-0.023	-0.023
			RMSE		
LSDV	0.187	0.125	0.072	0.048	0.029
Kiviet	0.059	0.044	0.030	0.023	0.017
Bun	0.061	0.045	0.030	0.023	0.017
GMM2	0.061	0.047	0.037	0.033	0.029
β=0.9			**Bias β**		
LSDV	-0.194	-0.129	-0.072	-0.046	-0.026
Kiviet	0.002	-0.001	-0.003	-0.002	-0.002
Bun	0.003	0.001	0.000	0.000	0.000
GMM2	-0.015	-0.020	-0.024	-0.025	-0.026
			RMSE		
LSDV	0.051	0.041	0.031	0.024	0.019
Kiviet	0.063	0.047	0.032	0.025	0.019
Bun	0.064	0.048	0.033	0.026	0.019
GMM2	0.064	0.048	0.033	0.026	0.020
β=0.8			**Bias β**		
LSDV	-0.216	-0.145	-0.082	-0.053	-0.030
Kiviet	0.001	-0.003	-0.004	-0.004	-0.003
Bun	0.003	0.001	-0.001	0.000	-0.001
GMM2	-0.019	-0.024	-0.029	-0.030	-0.031
			RMSE		
LSDV	0.054	0.044	0.035	0.028	0.023
Kiviet	0.069	0.052	0.037	0.029	0.023
Bun	0.071	0.053	0.037	0.030	0.023
GMM2	0.071	0.053	0.039	0.031	0.024
β=0.6			**Bias β**		
LSDV	-0.239	-0.164	-0.095	-0.065	-0.039
Kiviet	-0.002	-0.005	-0.007	-0.006	-0.005
Bun	0.002	0.000	-0.002	-0.001	-0.002
GMM2	-0.024	-0.029	-0.035	-0.038	-0.040
			RMSE		
LSDV	0.092	0.082	0.072	0.065	0.061
Kiviet	0.095	0.083	0.072	0.065	0.061
Bun	0.095	0.083	0.072	0.065	0.062
GMM2	0.095	0.104	0.091	0.066	0.062

Reference

Abramovitz, M. (1986), "Catching Up, Forging Ahead, and Falling Behind," *Journal of Economic History*, 46(2), 385–406.

Allerano, M. and S. Bond (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58(2), 277–297.

Anderson, T. W. and C. Hsiao (1981), "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, 76, 598–606.

Arbia, G. (2001), "The Role of Spatial Effects in the Empirical Analysis of Regional Concentration," *Journal of Geographical Systems*, 3(3), 271–281.

Barro, R. J. and X. Sala-i-Martin (1992), "Convergence," *Journal of Political Economy*, 100(2), 223–251.

Barro, R. J. and X. Sala-i-Martin (1995)a, "Technological Diffusion, Convergence and Growth," *Working Paper* 5151, National Bureau of Economic Research.

Barro, R. J. and X. Sala-i-Martin (1995)b, Economic growth (New York: McGraw-Hill).

Baumol, W. J. (1986), "Productivity Growth, Convergence and Welfare: What the Long-Run Data Show," *American Economic Review*, 1072–1085.

Blundel, R. and S. Bond (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87(1), 115–143.

Bun, M. G. B., and M. A. Carree (2005), "Bias-orrected Estimation in Dynamic Panel Models," *Journal of Business and Economic Statistics*, 23(2), 200–210.

Caselli, F., G. Esquivel and F. Lefort (1996), "Reopening the Convergence Debate: A New Look at Cross Country Growth Empirics," *Journal of Economic Growth*, 1(3), 363–389.

Cass, D. (1965), "Optimum Growth in an Aggregative Model of Capital Accumulation," *Review of Economic Studies*, 32, 233–240.

de la Fuente, A. (2002), "On the Source of Convergence: A Closer Look at the Spanish Regions," *Economica*, 70, 655–671.

Di Liberto, A., R. Mura and F. Pigliaru (2007), "How to Measure the Unobservable: A Panel Technique for the Analysis of TFP Convergence," *Oxford Economic Papers*, 1–26.

Doornik, J., M. Arellano and S. Bond (2002), "Panel Data Estimation Using DPD for Ox," mimeo, Institute for Fiscal Studies.

Friedman, M. (1992), "Do Old Fallacies Ever Die?," *Journal of Economic Literature*, 30(4), 2129–2132.

Gaduh, A. B. (2002), "Properties of Fixed Effects Dynamic Panel Data Estimators for a Typical Growth Dataset," *Economic Working Paper Series* 062, the Centre for Strategic and International Studies (CSIS: Indonesia).

Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge University Press (Econometrics Society Monograph, No. 11).

Hsueh, T. T. and Q. Li (1999), China's National Income, Westview Press, Boulder.

Islam, N. (1995), "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 111, 1127–1171.

Islam, N. (2000), "Small Sample Performance of Dynamic Panel Data Estimators in Estimating the Growth-Convergence Equation: A Monte Carlo Study," *Advances in Econometrics,* 15, 317–339.

Islam, N. (2003), "Productivity Dynamics in a Large Sample of Countries: A Panel Study," *Review of Income and Wealth*, 49(2), 247–272.

Islam, N. (2004) "What Have We Learnt from the Convergence Debate?" A. R. G. Donald, L. Oxley and K. I. Carlaw (eds), *Economic Growth: Theory and Empirics*, 83–135, Massachussets and Oxford, Blackwell Publishing Ltd.

Jian, T., J. D. Sachs and A. M. Warer (1996), "Trends in Regional Inequality in China" *Working Paper* 5412, National Bureau of Economic Research.

Judson R. and A. Owen (1999), "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists," *Economic Letters*, 65, 9–15.

Kiviet, J. F. (1995), "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models," *Journal of Econometrics*, 68, 53–78.

Koopmans, T. C. (1965), "On the Concept of Optimal Economic Growth," in *The Economic Approach to Development Planning*, Political Academy of Sciences, Amsterdam.

Lin, Y. (2001), *China's Regional Income Gaps: Analysis from Industrial Structure, Demography, and Education*, Nihon-Keizai-Hyoron-Sha (written in Japanese).

Mankiw, N. G., D. Romer and D. N. Weil (1992), "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407–437.

Minami, R. (1994), *The Economic Development of China*, MacMillan.

Nickell, S. (1981), "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49(6), 1417–1426.

Quah, D. (1993), "Empirical Cross-Section Dynamics in Economic Growth," *European Economic Review*, 37, 613–622.

Raiser, M. (1998), "Subsidising Inequality: Economic Reforms, Fiscal Transfers and Convergence Across Chinese Provinces," *Journal* *of Development Studies*, 34, 1–26.

Solow, R. M. (1956), "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70, 65–94.

Summers, R. and A. Heston (1991), "The Penn World Table (Mark 5): An Expanding Set of International Comparison. 1950–1988," *Quarterly Journal of Economics*, 106(2), 327–368.

Swan T. W. (1956), "Economic Growth and Capital Accumulation," *Economic Record*, 32, 334–361.

Weeks, M. and J. Y. Yao (2003), "Provincial Income Convergence in China, 1953–1997: A Panel Data Approach," *Econometric Reviews*, 22(1), 59–77.

Yao, S. (2001), "Regional Growth in China under Economic Reforms," *Journal of Development Studies*, 38(2), Dec. 167–186.

Young, A. T., M. J. Higgins and D. Levy (2007), "Sigma Convergence Versus Beta Convergence: Evidence from U.S. County-Level Data," *Emory Law and Economics Research Paper*, No. 07–4.