## Abstract

China's accelerated growth rate during the reform period 1978-97 has reinforced concerns about how to cope with continued expansion while also maintaining balanced regional growth. We examine both the tendency and the speed towards income convergence in among China's provinces during the two periods: the pre-reform period 1953-1977 and the reform period 1978-1997. The Solow growth model is used to provide the main theoretical framework. The empirical method accounts for heterogeneity in both initial technology and the rate of technological progress. Estimation problems including weak instruments and finite sample bias are addressed by the System GMM Estimator and a coefficient bound provided by the OLS and within group estimator, respectively. Although we find evidence of conditional convergence for both the periods, relative to the estimated convergence speed for other regions and countries, China's provincial convergence speeds are surprisingly low: 0.414% for the pre-reform period and 2.23% for the reform period. This means that along the course of the remarkable economic growth of China during the reform period 1978-97 the provincial income convergence process has been disappointing.

*Key Words and Phrases:* Provincial convergence, China, panel data, GMM estimators.

*JEL Classification:* O40, O41, O53, C14, H4

PROVINCIAL INCOME CONVERGENCE IN CHINA, 1953-1997: A PANEL DATA APPROACH

Yao Yudong and Melvyn Weeks[1,2]

October 2000

> *As China's economic miracle continues to leave millions behind, more and more Chinese are expressing anger over the economic disparities between the flourishing provinces of China's coastal plain and the impoverished inland, where 70 million to 80 million people cannot feed or clothe themselves and hundreds of millions of others are only spectators to China's economic transformation.*[3]
>
> *The New York Times*, 27 December 1995, p. 1.

# 1 Introduction

China's accelerated growth rate during the reform period 1978-97 has reinforced concerns about how to cope with continued growth whilst maintaining balanced regional income inequality. In September 1995 the Chinese government endorsed the view that regional inequalities have widened stating that "since the adoption of reforms and open door policies, we have encouraged some regions to develop faster and get richer, and advocated that the richer should act as a model for and help the poor. But for some reasons, regional economic inequalities have widened."[4]

In this paper we investigate both the tendency and the speed towards income convergence across the provinces of China during both the pre-reform period 1953-1977 and the reform period 1978-1997. The Chinese case is interesting for several reasons. First, China is the most populous country in the world and displays vast geographical disparities in the resource base, living standards and other determinants of income growth. Subsequently, these disparities place the issue of income convergence into sharp relief. Second, China has been subject to two contrasting policy regimes since the establishment of the P. R. China in 1949. Before the reforms of 1978 China was a closed economy under centralized planning. In 1978 China adopted economic reforms and the open-door policy, and as a result has experienced spectacular growth rates, enabling per capita incomes to almost quadruple during the last twenty years. In this context and faced with two distinct regimes characterised by a very different mix of market and non-market forces, we are

[1] Support from the Economic and Social Research Council under grant L138251008 is gratefully acknowledged.

[2] Faculty of Economics and Politics, University of Cambridge

[3] This quote is taken from Ravallion and Jalan (1996).

[4] People's Daily Overseas Edition, Oct. 5, 1995, p4

provided with a valuable opportunity to determine the impact of changes in market forces upon convergence speed. Finally, historical experience suggests that few countries have succeeded in maintaining political stability under conditions of severe income disparity. China's own history is full of uprisings, rebellions and revolutions sparked by economic injustice. As such the extent of regional income inequality is of considerable interest as it bears directly on the sustainability of economic reform and open-door policy.

Nonetheless, the question of whether against a backdrop of the removal of state controls of prices, and capital and labour allocations, inter-provincial income disparities have narrowed has been relatively neglected.

There have been three main studies on China's provincial income convergence: Jian, Sachs and Warner (1995), Chen and Fleisher (1996), and Raiser (1998). All studies have utilised the Solow growth model to examine income convergence. The study by Jian, Sachs and Warner (1996), using the cross-sectional approach, examines the convergence of GDP per capita across China's provinces[5] over the period 1952-1993, finding that real income convergence has been a relatively recent phenomenon, emerging strongly only since the reform period began in 1978. Chen and Fleisher (1996) also use the cross-sectional method finding evidence of both unconditional and conditional convergence of per capita income from 1978 to 1993. An estimated unconditional convergence speed of 0.9% per year implies that it would take 77.3 years to eliminate half of the gap in 1978. The conditional convergence speed (conditional on physical investment share, employment growth, human-capital investment, foreign direct investment and coastal location) is estimated to be 5.7% per year. This is considerably higher than the figure of 2% per year which is a relatively robust result from a number of *cross-region* studies (see, for example, Sala-i-Martin (1996), and Mankiw, Romer and Weil (1992) - hereafter MRW). A recent study by Raiser (1998) tests both unconditional and conditional convergence using both the cross-sectional and static panel data model during the reform period 1978-1992. He confirms previous findings that speed of the convergence falls over the course of economic reforms, with estimated convergence rates ranging from 0.8% to 4.2% per year for various sub-periods. Overall, previous studies suggest that there is no obvious convergence during the pre-reform period, with some consensus regarding conditional convergence during the reform period. However, there remains considerable uncertainty as to the speed of convergence.

The present study represents a significant contribution in a number of important aspects. First, there are two principal sources of bias in previous cross-sectional studies: unobserved province-specific heterogeneity in initial technology levels and endogenous explanatory variables. In particular, previous cross-sectional studies have assumed a homogeneous aggregate production functions across provinces and in doing so ignore the problem of heterogeneity in initial technology levels. This has considerable implications with regards the estimation of unbiased convergence rates. The problem of potential

[5] Only fifteen provinces in their study have GDP data for the pre-reform period 1952-77.

endogeneity of right-hand-side variables such as the investment rate is also neglected. By utilising both cross-sectional and time series variability, the panel data approach is seen as a promising alternative. In this study we use a first differenced system generalised method of moments estimator system (GMM) which allows us to eliminate unobserved country-specific initial technology effects and correct for endogeneity.

Finally, this study uses a more recent and extensive data source, namely the recently-released Hsueh-Li (1999) National Income data set, spanning 45 years from 1952 to 1995. Until 1993 National Income[6] (NI) had been the single most important economic indicator in the national and provincial accounts in China, and had occupied a position analogous to that of GDP in market economies. Data on NI are available in complete time series both at the national and provincial levels throughout the period 1952-1992. However, GDP data for more than half of China's provinces is only available after 1978, resulting in difficulties when comparing performance between the pre-reform and the reform period.

This paper is organized as follows. In Section 2 we examine the characteristics of the Chinese statistical reporting system, pre and post reform, and highlight the nature of the data used in this study. In Section 3 we motivate the ensuing analysis by presenting a brief graphical overview of the relative growth paths of China's provinces over the period 1952 to 1995. In Section 4 we present the Solow neoclassical growth model with a focus on its dynamics. Section 5 reviews the existing cross-sectional literature of testing income convergence. Section 6 introduces the panel data framework and System-GMM estimator employed in this paper. Section 7 contains the empirical results of using both the cross-sectional method based on OLS regressions and the panel data method. Section 8 concludes.

# 2 The Chinese Statistical Reporting System: Pre and Post Reform

Since the adoption of the comprehensive reform policy in 1978, China's statistical system has followed the reform wave. Starting from 1985, the state statistical bureau (SSB) has received financial aid from both the World Bank and the Asian Development Bank for further developing the national income accounts system, particularly for the estimation of GDP and its components in the system of national accounts (SNA) and to make up the deficiency of material product balances (MPS). SNA was completely implemented for the entire economy in 1992.

The Chinese statistical reporting system is organized within a hierarchical framework, from the top level, SSB, down to the provincial and county level. This organizational

[6]National Income (NI) is Net Material Product, which excludes depreciation of capital and the value added of the service sector. It includes value added of five "materials production": agriculture, industry, construction, transport, communication and commerce. NI differs from national income in the market economy in the coverage of the service sector.

structure forms a reporting system for the grass-roots units to submit statistical data to the statistical authorities. During the pre-reform period of 1952-77, the Chinese statistical system adopted the system of MPS. The making of MPS is tailored to meet the needs of the central planning economy. The main aggregate indicators are the value of total output and national income. National income is calculated as value-added and comparable with GDP in the SNA adopted by the market-economy countries. The main difference between GDP and National income accounts is that the latter do not account for the value-added of the service sector.

There are three principal shortcomings with the SNA. First, there is no provincial GDP data for about half of China's provinces during the pre-reform period. As a result, previous studies on China's regional economy have used national income data. Second, there are no provincial GDP deflators and the provincial retail price index is used to deflate NI for each province. Finally, there is no national and provincial investment deflators until 1991. With the support from SSB, Hsueh and Li (1999) have published the most complete set of Chinese national and provincial income accounts. The Hsueh-Li data set provides information for the service sector omitted in the material product balances and thus is directly comparable with those of other developed economies. Furthermore, provincial level GDP, GDP deflators, total investment, and investment deflators for both the reform and pre-reform periods, is also available. In this study we use provincial GDP data from the Hsuech-Li national income data set, covering the period 1952-95. Provincial GDP data and investment data for 1996 and 1997 is drawn from China Statistics Yearbook 1997.

Our data covers a 45 year period (1953-1997) and includes 28 provinces in China (see Table 1). The two exclusions are Hainan, a newly established province and Tibet, which does not have data until 1987. Nominal provincial GDP data is deflated using provincial GDP deflators. Real investment/GDP ratio is defined as the total investment within one year at 1995 constant price.[7] Because of the difficulty in obtaining data on capital depreciation ($\delta$), we assume that $\delta$ does not vary across provinces. This present paper maintains the assumption of Raiser (1998) that $g + \delta$ is equal to 0.07 and constant across China's provinces. Following Islam (1995), we use five-year time intervals for two periods: 1953-1978 and 1978-1997. By adopting this approach are results are less likely to be influenced by business cycle fluctuations.[8]

# 3 A Graphical Overview

Before we examine graphical evidence on differences between province level growth profiles, we say a few words regarding the choice of the Solow model as a framework for analysing convergence. Bernard and Durlauf (1996) highlight the very real differences

[7]The deflators of both GDP and investment take the prices of 1995 as 100.

[8]For 1952-1978 there are five sub-periods: 1953-57, 1958-62, 1963-67, 1968-72, and 1973-77. For 1978-1997 there are four: 1978-82, 1983-87, 1988-92, and 1993-97.

in the cross-section and time series approaches to testing for convergence. However, the critical point is that these tests are appropriate under very different circumstances. For example, whereas standard cross-section tests assumes that economies are far from a steady state, the time series tests focus on the *limiting* value of the difference between per capita output, and in this respect the emphasis is placed upon the equivalence of these limiting values. Subsequently, although the *sin qua non* of the cross-sectional approach is that initial conditions matter, the time-series approach focuses upon limiting distributions and thereby presupposes that initial conditions have no statistically significant effect on output differences. In this respect the time series approach, which in testing for convergence requires that economies are close to their steady state[9] is perhaps not such an appropriate framework to examine the China's provincial growth profiles which we believe reflect both initial conditions and the dynamics of a transition process.

In the closed economy Solow model the principal mechanisms behind any tendency towards convergence are factor mobility and technology flows. In this context and under ceteris paribus conditions we might expect a greater tendency towards convergence across *regions* given the existence of a common central government, institutions and a single currency. During the pre-reform period, the Chinese government had bothstrategic and ideological reasons for promoting the development of local autarchic economies and developing duplicate sets of industries in each province.

In addition, one of the distinguishing characteristics of China's provincial economy is autarchic, relatively closed local economies. In this particular respect the Solow growth model represents a suitable framework for testing convergence. Finally, we also emphasise that the objective of this paper is *not* to evaluate the respective merits of alternate frameworks with which to examine convergence, but more simply to examine the implications of using a cross-sectional compared to a panel data approach for the analysis of convergence within the confines of the Solow growth model. In this respect our approach is similar to Nerlove (1999).

The study of provincial convergence in China is of particular interest given the opportunity afforded to determine the extent to which changes in government determined factor mobility have affected convergence.[10] During the pre-reform period China was a centrally-planned economy and internal migration of labor was controlled by the government.[11] However, in the reform period it has become possible to have unregistered migration, although the constraints upon unregistered migration across provinces are still powerful.[12] Similarly, there is no significant evidence for inter-provincial capital mobility.

[9] This follows since observed sample moments are used to approximate the limiting moments.

[10] Obviously without a valid counterfactual we cannot isolate the direct effect of a set of policy reforms such as the removal of restrictions on labour mobility.

[11] The principal instrument for such control lies in the 1958 Regulation on the Registration of Households, which requires every household to register its place of residence, and to gain permission for any changes in residence. Until the early 1980s, migration without permission was extremely difficult.

[12] The main reason is that illegal migrants lack access to schooling for children, state-run health care and other services.

For the pre-reform period, state-sector investment (accounting for almost all investment across provinces) was centrally allocated. In contrast, during the reform period, a number of studies have argued that capital mobility in China as a whole has declined (see, for example, World Bank Country Study (1994)). Young (2000) notes that the process of decentralization resulted in the fragmentation of the domestic market and interregional barriers to trade.

Below we provide a graphical analysis of the provincial growth process during the pre-reform (1952-77) and reform periods (1978-97). Our intention is to present some sense as to the general trend toward income convergence[13] or divergence among China's 28 provinces. We first categorise provinces on the basis of coastal (9 provinces) and the interior provinces (19 provinces). One reason for adopting this classification is that there is a significant growth difference between the coastal provinces and the interior provinces, in part attributable to regional policy during the reform period which has emphasized the development of the coastal region. In Table 1 we list all the provinces and indicate location in terms of the coastal/interior classification. Figure 1 shows that the income gap in terms of log GDP per capita between coastal and interior provinces increases slightly after 1978. A problem with this classification of provinces is that there exists considerable heterogeneity *within* both the coastal and interior provinces. For example, not all coastal provinces outperform the interior provinces in terms of the long-run growth performance. In fact only the five fastest growing provinces (Guangdong, Fujian, Jiangsu, Shandong and Zhejiang), have outstanding growth performance relative to the interior provinces. In addition the annual average growth rates of three coastal provinces (Tianjiang, Liaoning and Shanghai) are less than the national average (see Table 2). Figure 2 shows that the average per capita income of the five fastest growing coastal provinces[14] is rapidly approaching that of Shanghai. In contrast, two provinces, Heilongjiang, the heavy industrial base in North-east of China, and Qinghai, an inland province in North-west of China, have been obviously falling behind. Figure 3 compares the remaining coastal provinces (Hebei, Liaoning, Tianjing and Shanghai itself) with the interior provinces excluding Heilongjiang and Qinghai.

We also divide 9 coastal provinces into the richest (Shanghai), the average group and the poorest (Hebei). Figure 4 demonstrates that the income gap between the richest coastal province (Shanghai) and the poorest coastal (Hebei) is relatively constant. Similarly, we divide 19 interior provinces into the richest (Beijing), the average group and the poorest (Guizhou). Figure 5 shows that the income gap between the richest interior province (Beijing) and the poorest (Guizhou) has increased slightly since 1980s.

Based upon the graphical evidence presented above there is little evidence of overall income convergence among all provinces during both the pre-reform and reform period. Furthermore, as all graphs show the income levels using logarithmic scales, the slopes

[13]This analysis considers convergence as the reduction of inter-province income differences over time.

[14]These provinces have grown at more than 10% each year since 1978. These provinces have also been referred to as the 'five dragons'.

represent per capita growth rates. Therefore, since the growth rates of all provinces appear to be moderately steady since 1978, this implies that the significant initial differences between provincial growth rates is maintained.

# 4 Solow Neoclassical Growth Framework

Despite its age the Solow growth model continues to provide the theoretical basis for a large number of the cross-sectional studies of income convergence (see, for example, Barro and Sala-i-Martin (1991), and MRW). More recently, studies by Islam (1995), CEL (1996), Lee Pesaran and Smith (1997, LPS), and Nerlove (1998), have utilised both cross-section and time series observations to test income convergence and estimate the convergence rate. Since the estimated convergence speed has important policy implications, one of the critical issues is how to consistently estimate this parameter. To facilitate comparison between cross-sectional and the panel data studies, we provide a brief overview of the Solow growth model.

Using standard notation we assume a Cobb-Douglas production function with output ($Y$) and three inputs, capital ($K$), labour ($L$) and labour augmenting technological progress ($A$).[15] Assuming constant returns to scale we write

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

where $0 < \alpha < 1$. Labor force and technology grow at the following constant and exogenous rates

$$L(t) = L(0)e^{nt}, \tag{2}$$

$$A(t) = A(0)e^{gt}, \tag{3}$$

where $n$ is the growth rate of the labor force and $g$ is the rate of technological progress. $L(0)$ is the initial state of the labor force and $A(0)$ is the initial state of technology. If we denote $\widehat{y}(t) = \frac{Y(t)}{A(t)L(t)}$ as output per effective unit of labor, and $\widehat{k}(t) = \frac{K(t)}{A(t)L(t)}$ as capital per effective unit of labor then $\widehat{y}(t) = f(\widehat{k}(t)) = \widehat{k}(t)^{\alpha}$. The evolution of capital is given by

$$\dot{\widehat{k}}(t)\ s\widehat{k}^{\alpha}(t) - \widehat{k}(t)(n+g+\delta), \tag{4}$$

where $s$ is the savings rate and $\delta$ is the depreciation rate of capital. Capital is subject to diminishing marginal returns such that $\frac{df(\widehat{k}(t))}{d\widehat{k}(t)} > 0$ and $\frac{d^2 f(\widehat{k}(t))}{d\widehat{k}(t)^2} < 0$. The steady state capital stock, $\widehat{k}^*$, can be determined by setting equation (4) equal to zero

$$\hat{k}^*(t) = \left(\frac{s}{n+g+\delta}\right)^{\frac{1}{1-\alpha}}. \tag{5}$$

[15] The exposition of the Solow model is mainly based on MRW and Islam (1995).

By substituting (5) into the production function $\widehat{y}(t) = \widehat{k}(t)^{\alpha}$, the steady state output per effective unit of labor can be derived, which in logarithmic form may be written

$$\ln \widehat{y}^{*} = \left(\frac{\alpha}{1-\alpha}\right)[\ln s - \ln(n+g+\delta)]. \tag{6}$$

Following Islam (1995), approximating around the steady state, the rate of convergence, $\lambda$, denoting the rate at which output per effective unit of labor approaches its steady state value, is given by

$$\frac{d \ln \widehat{y}(t)}{dt} = \lambda[\ln(\widehat{y}^{*}) - \ln \widehat{y}(t)], \tag{7}$$

where $\lambda = (1-\alpha)(n+g+\delta)$. The solution to the differential equation (7) is

$$\ln \widehat{y}(t_2) = (1-\zeta) \ln \widehat{y}^{*} + (1-\zeta) \ln \widehat{y}(t_1), \tag{8}$$

where $\zeta = e^{-\lambda\tau}$, and $\tau = (t_2 - t_1)$. This equation represents a partial adjustment process where the optimal or target value of the dependent variable is determined by the explanatory variables of the current period. In the present case, $\widehat{y}^{*}$ is determined by $s$ and $n$, which are assumed to be constant for the entire intervening time period between $t_1$ and $t_2$ and hence represent the values for the current year as well.

In operationalising this model we need to work with output *per capita*. In order to reformulate equation (8) we express output per effective unit of labor as

$$\ln \widehat{y}(t) = \ln\left(\frac{Y(t)}{A(t)L(t)}\right) = \ln\left(\frac{Y(t)}{A(t)e^{gt}L(t)}\right),$$

or

$$\ln \widehat{y}(t) = \ln\left(\frac{Y(t)}{L(t)}\right) - \ln A(0) - gt. \tag{9}$$

Substituting the expression for $\ln \widehat{y}(t)$ into equation (8) and subtracting $\ln y\,(t_1)$ from both sides, provides us with an expression of per capita output growth over the period $t_2 - t_1$

$$\begin{aligned} \ln y\,(t_2) - \ln y\,(t_1) \;=\; & -(1-\zeta)\ln y(t_1) \\ & +(1-\zeta)\ln A(0) + g(t_2 - \zeta t_1) \\ & +\underbrace{(1-\zeta)\frac{\alpha}{1-\alpha}\ln(s) - (1-\zeta)\frac{\alpha}{1-\alpha}\ln(n+g+\delta)}_{z}, \end{aligned} \tag{10}$$

where $y(t) = \frac{Y(t)}{L(t)}$ stands for per capita output, and the parenthetical term $z$ denotes the log of the steady-state per capita output. The Solow model in equation (10) focuses upon the transitional growth dynamics of one economy to its steady state income path. Letting $\beta = -(1-\zeta)$ denote the parameter on income at $t_1$, the speed of convergence is then given by

$$\lambda = -\frac{\ln(\beta+1)}{\tau}.$$

Below we examine identification conditions for both the OLS cross-sectional and panel data approach. In this discussion a key issue will be the interpretation and measurement of $A(0)$.

# 5 An Overview of Cross-sectional Studies

Much of the literature on cross-sectional convergence focussed upon testing two important concepts: $\beta$ unconditional and $\beta$ conditional convergence. Within the Solow growth framework, $\beta$ *unconditional* convergence implies that the rate of savings, population growth, technological progress, and depreciation ($s$, $n$, $g$, and $\delta$, respectively) and initial technology, $A(0)$, are *equal* across countries; in the context of a cross-sectional study all these terms become constant. To see this we first write equation (10) as

$$g_{i,t_{21}} = a + \beta y_{i,1} + u_i, \tag{11}$$

where $i = 1, ..., N$ denotes the region index, $g_{i,t_{21}} = \ln y_i(t_2) - \ln y_i(t_1)$ denotes the growth rate of real per capita GDP for region $i$ over the time period $t_2 - t_1$, and $y_{i,1}$ is the initial level of real per capita GDP at $t_1$. $\beta < 0$ in (11) implies $\beta$ convergence, in that growth depends (negatively) only on the initial level of per capita income.

Of particular interest is the intercept term $a$. Within the context of an unconditional model of convergence, and based upon (10), $a$ is given by

$$\begin{aligned} a &= (1-\zeta)\ln A(0) + g(t_2 - \zeta t_1) \\ &\quad +(1-\zeta)\frac{\alpha}{1-\alpha}\ln(s) - (1-\zeta)\frac{\alpha}{1-\alpha}\ln(n+g+\delta). \end{aligned} \tag{12}$$

Here $a$ is additive in the effect of savings, population growth, initial technology, $A(0)$, and the exogenous rate of technological change, $g(.)$. Different initial conditions affect the speed of convergence to the common steady state. As both MRW and Islam (1995) note, if we write $\ln A(0) = \varpi + \varepsilon_i$, namely as additive in a constant and an unobserved region specific term, then $u_i$ in (11) will have an error components representation given by $u_i = \epsilon + \varepsilon_i$, where $\epsilon$ is a random disturbance term. For any given region the unobserved region specific effect $\varepsilon_i$ is likely to be correlated with savings and population.

Earlier cross-sectional studies (e.g. Baumol 1986) found that $\beta$ unconditional convergence has taken place only among relatively homogeneous groups of countries such as members of the OECD, but not among a broader number including developing countries. Studies by Barro and Sala-i-Martin (1991, 1992, 1995) and Sala-i-Martin (1996) have examined the tendency of regions *within* a country to exhibit income convergence. For example, Sala-i-Martin (1996) argues that $\beta$ convergence occurs within a number of countries including US states, Japanese and European countries. Furthermore, the speeds at which the regions of different countries converge over different time periods are surprisingly similar (about 2%), which implies that 50% percent of this difference vanishes only after 115 years.

The concept of $\beta$ *conditional* convergence represents an extension to the unconditional model, and by controlling for differences in population growth rates, and the savings rate, removes the necessity to assume identical steady state levels of income for all countries. We now rewrite (11) as

$$g_{i,t_{21}} = a + \beta y_{i,1} + \boldsymbol{\theta}'\mathbf{x}_i + u_i, \tag{13}$$

where $\mathbf{x}_i$ is a vector including $\ln(s_i)$ and $\ln(n_i + g + \delta)$, and $\boldsymbol{\theta}$ is the associated vector of coefficients. As a result the intercept $a$ may now be written as $a = (1-\zeta)\ln A(0) + g(t_2 - \zeta t_1)$. However, given that parameter identification is based solely upon cross-sectional information, it is still necessary to enforce homogeneity of both $\ln A(0)$ and $g$. In this respect a finding of conditional convergence implies that the economies converge to the *same* underlying steady-state path.

MRW concluded that there is strong evidence for $\beta$ conditional convergence among a broad number of countries over the period 1960-1985, once saving, population growth and human capital accumulation are controlled for. Estimates of the average speed of convergence range from 0.6 to 2 percent per annum.[16] However, Islam (1995) and CEL (1996) question the validity of this finding, stating that problems with the cross-sectional approach, and in particular the bias generated by the correlation of omitted *country specific* technology effects and the steady state technology growth rate, $g$, cause the speed of convergence parameter to be biased downwards. The essence of their argument is that studies such as Barro (1991) and MRW are based on the assumption that the initial state of technology, $A(0)$, and $g$ are homogeneous across countries. As such, the estimated speed of convergence gives the average speed at which a country converges to the *common* steady state income path. If the homogeneity assumption of $A(0)$ and/or $g$ is not valid, convergence estimates will be biased (see, for example, LPS). Further, there may be additional problems due to endogeneity of control variables. For example, Cho (1996), finds endogeneity of both the investment rate and the population growth rate in the Summers-Heston data set.

The panel data approach initiated by Islam (1995) and CEL (1996) has potential advantages over the cross-sectional method in overcoming these two problems. However, whether the potential advantages can be realized largely depends on the panel data estimators used, and in the case of an endogeneity correction, the availability of feasible instruments. We examine these issues below.

# 6 Empirical Method of Panel Data

## 6.1 Panel Data Framework

If we collect terms involving the log of the first period GDP on the right-hand side of equation (10), we may write an autoregressive form of the growth model as

$$\begin{aligned}\ln y(t_2) = {} & \zeta \ln y(t_1) \\ & +(1-\zeta)\ln A(0) + g(t_2 - \zeta t_1) \\ & +(1-\zeta)\frac{\alpha}{1-\alpha}\ln(s) - (1-\zeta)\frac{\alpha}{1-\alpha}\ln(n+g+\delta).\end{aligned} \tag{14}$$

[16] This finding implies that it takes between 35 and 115 years to close half the gap between initial income and the common steady state income.

(14) represents a general dynamic framework within which to examine income convergence. Using obvious notation we may, adding a disturbance term, write this equation as

$$y_{i,t} = by_{i,t-1} + \boldsymbol{\theta}'\mathbf{x}_{i,t} + T_t + \eta_i + v_{i,t}, \tag{15}$$

where $\mathbf{x}_{it} = (\ln(s_{it}), \ln(n_{it}+g+\delta))'$, $\boldsymbol{\theta} = ((1-\zeta)\frac{\alpha}{1-\alpha}, -(1-\zeta)\frac{\alpha}{1-\alpha})'$ and $b = 1+\beta = \zeta$. We interpret the effects $\eta_i$ as a composite of unobservable province-specific factors, thereby representing the combined effect of institutions, factor endowments, and relative location, together with initial technology differences. Similarly, the $T_t$ captures the time-specific effects which include the rate of technological change.

In comparing (15) with (13) we note that the intercept in (13) is additive in two constant terms: $(1-\zeta)\ln A(0)$ and $g(t_2 - \zeta t_1)$. With multiple observations per cross-sectional unit it is possible, through the introduction of province-specific and time-specific effects, to relax the assumption of strict parametric homogeneity of both $\ln A(0)$, the initial technology state, and $g$, the technological progress rate. Islam (1995) and CEL (1996) allow the aggregate production function to differ across countries with respect to initial technology state $A(0)$ while the homogeneity of $g$ is maintained.

By allowing for differences in the initial state of technology $A(0)$ we thereby allow economies to have distinct but parallel steady state income paths. However, the economic implications of the convergence parameter become weaker. As Durlauf and Quah (1998) note, although the use of panel data methods facilitates a control for unobserved heterogeneity, there are subsequent difficulties in interpreting the output from convergence regressions. In particular, if we allow $A(0)$ to differ across economies, a finding of convergence to the then parallel underlying steady-state paths is not informative in the sense that we are not able to make statements as to whether poor economies are converging to the (common) steady state. Nevertheless, and conditional upon a homogeneous growth rate of technology, we can say that the income gap among the economies is stable due to the existence of parallel steady-state growth paths. In this respect convergence may be interpreted as the speed with which economies approach parallel long-run growth paths.

## 6.2 Heterogeneity in the Rate of Technology Progress Rate

The question whether or not economies have the same technology progress rate is also of considerable empirical interest. LPS (1997) demonstrate that if the assumption that the technological progress rate $g$ is the same across all countries or provinces is invalid, then the effect of imposing a common $g$ biases the estimate of $b$ toward 1[17]. Lee, Pesaran, and Smith (1998) suggest the use of the panel data approach to allow countries to differ not only with regard to $A(0)$ but also with respect to $g$. If $g$ does vary across the economies the long-run growth paths are not parallel and the system of economies are diverging. As far

[17] LPS (1997) demonstrate the existence of pervasive heterogeneity in $g_i$ in the Summers-Heston data.

as economic implications for income inequality is concerned, the existence of heterogeneity in technology growth rates completely alters the notion of (conditional or unconditional) convergence. As Lee, Pesaran, and Smith (1998), such a finding provides no insight into a comparison of the evolution of the distribution of cross-country outputs.

Although an assumption of homogeneity of the rate of technological progress across provinces during the pre-reform period is tenable, this is unlikely to be the case for the reform period. In particular, we might conjecture that the technology progress rate of coastal provinces is different from that of the interior provinces. This hypothesis can be tested empirically by the use of a composite dummy based upon taking the product of a time and coastal dummy. We do this by rewriting 15 as

$$y_{i,t} = by_{i,t-1} + \boldsymbol{\theta}\mathbf{x}_{it} + D_j T_t + T_t + \eta_i + v_{i,t}, \tag{16}$$

where $D_j = 1$ if the $ith$ province belongs to the coastal region. Note that within each region (coastal and interior) there is a maintained assumption that the rate of technological progress is homogeneous.

### 6.3 The System GMM Estimator and the Weak Instruments Problem

Estimates of the convergence speed parameter based upon panel data studies have been, in general, considerably higher than the 2% of the cross-sectional studies. For example, Islam (1995) using the Summers-Heston data finds annual convergence rates between 3.8% and 9.1%, whereas CEL using the first-difference GMM ( DIFF-GMM) estimator finds a convergence rate of 10%. However, the estimation problems of the panel data estimators used have been largely ignored[18]. Here we concentrate on CEL's first-difference GMM estimator. Since the true value of the autoregressive parameter, $b$, is very likely to be near 1 (as shown in Nerlove (1998)) the dependent variable $y_{i,t}$ is thus very persistent and nearly nonstationary. Blundell and Bond (1998) use Monte Carlo simulations to demonstrate the relatively large finite-sample bias generated by the use of the DIFF-GMM estimator to estimate autoregressive models from moderately short panels. Table 3 reports some of their results showing the poor performance of the DIFF-GMM estimator for high values of $b$. The DIFF-GMM estimator exhibits both a large downward bias and very imprecise estimates.

[18]The within-groups estimator used by Islam (1995) provides downward biased and inconsistent estimates of the autoregressive parameter in a dynamic panel model with fixed $T$ (see Nickell (1981)). Nerlove (1998) uses a random-effects maximum-likelihood estimator which is unconditional on the initial observations. However, we believe that the random-effect assumption may be invalid for three reasons. First, the random effect has to be uncorrelated with the exogenous variables included in the model such as investment rate. Islam (1995) obtains the estimates of the individual country effects and finds that the estimated country effects have a positive relationship with both the observed growth rates and human capital. Moreover, the distribution of these country effects is not normal. Finally, in the finite sample the individual country effects may be better taken as fixed effects term rather than a normal distribution.

The System GMM estimator developed by Blundell and Bond (1998) represents a significant improvement over the first difference GMM estimator proposed by Caselli, Esquivel and Lefort (1996) (hereafter CEL). Given that annual log per capita GDP (or five-year averages of log per capita GDP), is likely to be persistent, weak correlations exist between the growth rate of log per capita GDP and the lagged log per capita GDP levels. This results in what is referred to as a *weak instruments* problem in the context of the first-difference GMM estimator. Blundell and Bond (1998) demonstrate that this problem can result in large finite-sample biases when using the first-difference GMM estimator to estimate autoregressive models for moderately persistent series from relatively short panels. The authors show that the biases can be dramatically reduced by incorporating more informative moment conditions that are valid under quite reasonable stationarity restrictions on the initial condition process. Essentially the System GMM estimator represents the use of lagged first-differences as instruments for equations in levels, in addition to the usual lagged levels as instruments for equations in first-differences. In addition, the finite-sample performance of the System GMM can be tested by the identification of an estimation range for the convergence speed provided by the OLS and within-group estimator.

As a way of correcting for this bias Blundell and Bond (1998) note that if the initial conditions $y_{i1}$ satisfy the stationarity restriction $E[\Delta y_{i2}\eta_i] = 0$, then suitably lagged values of $\Delta y_{i,t-s}$ as well as $\Delta x_{i,t-s}$ are available as instruments for the levels equations. The resulting System GMM estimator combining equations in levels with equations in first-differences, is shown to provide dramatic gains both in asymptotic efficiency and in finite sample properties, relative to the first-difference GMM estimator (see Arellano and Bond (1991)). Both sets of moment conditions can be exploited as a linear GMM estimator in a system context.

To evaluate the finite sample performance of System-GMM estimator in this particular application we establish a bound for the autoregressive parameter. Knowledge of the direction of the bias in the OLS and within-groups estimator in models with fixed T is used to establish an approximate upper and lower bound for $b$ in the growth regressions. As Hsiao (1986) shows, omitting unobserved time invariant country effects in a dynamic panel data model will cause OLS levels estimates to be biased upwards and inconsistent, given the positive correlation between the lagged dependent variable $y_{i,t-1}$ and the permanent effects $\eta_i$. On the other hand, the within-group estimator produces a downward bias with the extent of attenuation increasing when exogenous covariates are added. Sevestre and Trognon (1996) demonstrate that, for the purely autoregressive case (i.e. $\boldsymbol{\theta} = \mathbf{0}$) and within a class of estimators including the mean group estimator and the Generalized Least Squares estimator, the within group and OLS estimator can provide a tighter bound around $b$. Nerlove (1998) empirically confirms that the Sevesre-Trognon bound generally holds except for one reversal in the case of the first-difference model.

# 7 Results

Table 4 reports the results for unconditional $\beta$ convergence over the pre-reform period 1953-1977 and the reform period 1978-1997. Focussing upon parameter estimates in columns 1 and 3, we observe that the coefficients on the log of initial per capita GDP in 1953 and 1978 are negative but insignificant for both periods, and thus we find no tendency for poor provinces to grow faster on average than rich provinces. In terms of the pre-reform period, one possible explanation for this finding is the existence of two large negative shocks (the Great Leap Forward 1958-61 and the inception of the Cultural Revolution 1966-70). Since 1978, the economic reform has introduced market forces across all provinces and the overall economic growth process has been fairly steady. In this context the finding of no $\beta$ unconditional convergence for the reform period 1978-97 is surprising. These results stand in sharp contrast with the existing evidence on unconditional convergence where the cross-sectional units are regions within a single country. However, our results are similar to the findings of many previous studies on the failure of incomes *across countries* to converge (see, for example, Baumol (1986)). In this respect our findings suggest that China's provinces exhibit greater heterogeneity than regions within countries such as the U.S.

In columns 2 and 4 we introduce a coastal dummy which proxies for differences between the coastal and interior provinces in terms of the initial technology level, technology progress rate and the steady-state income. As a result it is necessary to be careful in interpreting our results given that conditional on a significant coastal dummy it is not possible to say whether the technology progress rate is different between the coastal and interior provinces. The coefficient of the coast dummy is not significant for the pre-reform period 1953-1977. The coefficient on the log of initial per capita GDP is negative but insignificant. In contrast the coefficient on the coast dummy is highly significant for the reform period, and in addition the coefficient on the log of initial per capita GDP is significantly negative. This is consistent with the previous findings such as Jian, Sachs and Warner (1995). Further, the inclusion of the coast dummy substantially improves the fit of the regression from 5.5% to 54.8%. Unlike the previous studies, however, we do not claim that there is clear evidence for $\beta$ convergence in the reform period 1978-97. The reason for this stems from the fact that it is possible that the significanceof the coastal dummy is driven by differences (between coastal and interior provinces) in the underlying rate of technological progress, thereby implying system-wide income divergence. In fact, this is an entirely plausible scenario since during the pre-reform period China was extremely isolated from the outside world such that geographical location had little bearing in accessing technical knowledge. After 1978 China has adopted an open door policy. Therefore, access to the ocean, and thereby to the world market, has undoubtedly afforded a real advantages to the coastal provinces. In the context of a fragmented domestic economy of China and restrictive interregional trade (see, for example,Young (2000)), the interior provinces may not be able to share the same technical knowledge as the coastal provinces.

Table 5 presents the cross-section results for $\beta$ conditional convergence for both periods. Control variables are the investment/GDP ratio ($I/Y$) and the population growth rate ($n + g + \delta$). As before we begin by examining parameter results in columns 1 and 3. For the pre-reform period the coefficient on the initial level of income is positive but insignificant; that is, there is no evidence of divergence or convergence. This finding is consistent with the results of Jian, Sachs and Warner (1996), namely that there is no strong conditional convergence across the provinces during the pre-reform period. For the reform period, the coefficient is negative but insignificant and in this respect contrary to the current consensus of conditional convergence across both countries and regions[19]. One possible reason is that we use a new provincial GDP data especially with province-specific GDP deflator and investment deflator relative to the previous studies. The finding of an insignificant coefficient of investment is consistent with the results of Chen and Fleisher (1996) and may imply that capital accumulation is not the main source of rapid economic growth of China after 1978.

After controlling for the investment rate and population growth, a significant coastal dummy may be interpreted in terms of diffferences in the initial technology level, and/or in the rate of technological progress. For both periods the coastal dummy variable is statistically significant. For the pre-reform period, the coastal dummy is more likely to reflect heterogeneity in the initial technology level. The coefficient on the initial level of income remains insignificant; that is, there is no evidence of $\beta$ conditional convergence. In the reform period the inclusion of a coastal dummy significantly improves the fit of the regression with $R^2$ increasing from 0.81% to 53.8% The coefficient on the initial level of income is significant, consistent with the findings of the previous studies. Again we note that this finding cannot be interpreted as evidence of $\beta$ conditional convergence given that the coastal dummy may reflect the heterogeneity of technology progress rate during the reform period. If so, there is evidence of a system-wide income divergence.

Finally, we cannot exclude the possibility that the reasons for these unexpected findings result from the underlying problems of the cross-sectional method. Namely, that cross-sectional studies fail to capture the dynamics of income convergence, the initial heterogeneity in $A(0)$ across the provinces, and the potential endogeneity problem of the investment/GDP ratio and the population growth rate. We explore this proposition below.

## 7.1 Panel Data Test for Conditional Convergence

In Table 6 reports results for the growth regressions for the pre-reform period based on equation (15)[20]. A positive coefficient less than 1 on initial GDP is interpreted as

[19]For example, MRW finds conditional convergence for different country groups such as non-oil countries during the period 1960-1985.

[20]All reported standard errors are corrected for heteroskedasticity. The parameter estimates and standard errors reported from GMM estimation are one-step estimators. All regressions except the first-difference GMM regressions include time dummies and they are found to be jointly significant in every

conditional convergence. Columns 1 and 2 report, respectively, the results of the OLS levels estimator and the within groups (WG) estimator. The OLS estimator provides a much higher estimate of the autoregressive parameter, $b$, than the WG estimator, together providing the upper and lower bounds for the GMM estimation. The coefficient of $\ln y_{t-1}$ is greater than 1 implying that provincial income is highly persistent. The implied speed of convergence based on the results of the WG estimator in column 2, 19.37%, is significantly different from the divergence implied by the OLS estimation.

In column 3 we present parameter estimates using the DIFF-GMM estimator, assuming that initial income, investment, together with population growth, are endogenous. The coefficient on the lagged dependent variable is statistically significant and exceeds the upper bound provided by within groups estimator. Given that $\ln y_{i,t}$ is highly persistent according to the OLS estimation in column 1, it is likely that the poor performance of the DIFF-GMM estimator is due to the weakness of the instrument set used. Furthermore, an application of the Sargan Test suggests that the instrumental variables used in the DIFF-GMM estimator are not valid.

Since the DIFF-GMM estimator appears to suffer from the weak instruments problem we use a system GMM estimator as an alternative. The results are reported in column 4. The coefficient on the lagged dependent variable from the system-GMM estimator falls between the upper and lower bound (0.3797 to 1.0248), and therefore is more likely to be unbiased. In addition, the estimates of the population growth rate and investment ratio are more precise than those obtained from first differenced GMM estimator. An application of the Sargan test (p-value=0.416) strongly suggests that the instrumental variables used in the system GMM are valid. Thus, based upon this evidence the system GMM is the preferred estimator for the Solow growth regression during the pre-reform period. The estimated speed of convergence is approximately 0.41% per year and extremely slow.

All regressors except the population growth rate are statistically significant at the 5% level. The coefficients of the investment ratio is negative and significant in the system GMM regression. During the pre-reform period, all provincial investment was controlled by the state and made through China's state-own enterprises. In this respect the negative sign of investment rate may be interpreted in the light of soft budget constraints,[21] which represented a major incentive problem of the state-owned enterprises during the central planning period.

Table 7 reports growth regressions for the reform period. Similar to the analysis for the pre-reform period the system GMM estimator is preferred to the DIFF-GMM estimator. First, the convergence coefficients of $\ln(y_{t-1})$ of the DIFF-GMM estimator (1.0299) falls outside of the lower-upper bound (0.9043-0.9927) given by the within group and OLS

regression. In order to conserve space the time dummy estimates are not reported in the tables. All parameter estimates are obtained using DPD98 for Gauss (see Arellano and Bond (1998)).

[21] Soft budget constraints are a key characteristic of socialist economies and remain an important concern in transition economies. Following Kornai (1980), a local government or an enterprise is said to have a soft budget constraint when it expects to be bailed out when faced with financial difficulties. This creates an obvious incentive problem.

estimator respectively, whereas the system GMM estimate falls within the bound. Second, the OLS estimate of the autoregressive parameter shows that provincial income is highly persistent. The implied convergence rate based upon the system GMM estimator is 2.23% per year. An application of the Sargan Test suggests that the instrumental variables used in both the DIFF-GMM system GMM estimator are significant.

The coefficient of investment has a positive sign across all estimators, although they are insignificant. In contrast to the pre-reform period, the positive investment coefficient may be due to the improvement in investment efficiency following the changes in China's investment ownership structure after 1978, with increasing shares of private investment and FDI. The process of increasing market competition and fiscal decentralization have hardened budget constraints of the state-owned enterprises (Qian and Roland (1998)). Nevertheless, the general insignificance of the coefficients of investment rates for the reform period 1978-97 may imply that inter-provincial differences in productivity instead of capital accumulation are likely to play an important role in explaining differences in provincial income. The coefficient on the population growth rate has the expected sign in all regressions.

#### 7.1.1 A Test for Heterogeneity in the Rate of Technological Progress

Whether the coastal provinces have the same rate of technological progress as the interior provinces is a critical concern for our study. Table 8 reports the system GMM estimates based upon equation (16)[22]. For the pre-reform period, using a Wald test, we cannot reject the null hypothesis that across interior and coastal provinces there is a homogeneous rate of technological progress (p-value=0.303%). The implied convergence speed is 0.879% per year (the half-life is 77 years) which although higher than the estimated convergence speed using (15) is still very slow. For the reform period the interaction of the two dummies is highly significant (p-value=0.00%), implying a difference in the technology progress rate of the coastal and interior provinces. An immediate implication from the Solow growth model is that the group of the long-run growth paths of the coastal provinces is not parallel with that of the interior provinces. As such the finding of conditional convergence ($b = 0.8467 < 1$) implies that the coastal and interior provinces are approaching their own parallel steady state paths at the same convergence speed (3.328% per year; the half-life is 22 year), and thus form two distinct convergence clubs. Our findings, namely that by imposing homogeneity in the rate of technological progress the estimate of $b$ is pushed towards one, and the measured speed of convergence is reduced, is consistent with the results obtained by LPS (1997).

The investment rate has a negative sign for the pre-reform period, which may be rationalized by the existence of a central planning regime and the pure state investment

[22] As with the results of the panel data tests with imposed homogeneity in rate of technological progress rate, the system GMM estimator is preferred in the growth regressions for both the pre-reform and reform period. To conserve space, we do not report the results of OLS, Within Group and the first-difference GMM estimator.

structure. For the reform period, it has a positive sign which may be interpredted in the light of increasing private investment and FDI. Nevertheless, the population growth and investment rate are not significant for both periods.

# 8 Conclusion and Qualification

This paper extends the analysis of previous studies on China's provincial income convergence by using the System-GMM estimator to examine provincial income convergence tendency and speed during the pre-reform period 1953-77 and the reform period 1978-97.

We summarize our empirical findings:

1. Using the cross-sectional approach as a benchmark for the panel data studies our findings are striking. For both the pre and post-reform period we find no evidence of $\beta$ unconditional or conditional convergence. This result echoes the previous findings of Jian, Sachs and Warner (1996).

3. Using the DIFF-GMM estimator we find that most coefficients fall outside the defined bounds given by the OLS and within groups estimator. A system GMM estimator is shown to be the preferred estimation method.

2. Using the panel data approach we find evidence of provincial income convergence. These relative findings are explainable in terms of a correction for both the homogeneity and exogeneity assumptions which characterise the single cross-section regression. From the perspective of growth theory these results highlight the significance of the differences in the aggregate production functions across China's provinces. Moreover, the obtained convergence speeds are 0.414% for the pre-reform period and 2.23% for the reform period using the system GMM estimator. They are very low relative to the panel data findings for the country studies (Islam, 1995; CEL, 1996).

3. Homogeneity in the rate of technological progress rate is a maintained hypothesis of the panel data results noted above. Since the beginning of the reform period the coastal provinces' growth performance has been much more remarkable than the interior provinces of China. Based on the empirical specification of LPS (1998), we allow for heterogeneity in technology growth rate between the coastal and interior provinces by using a coastal dummy interacted with the time dummy. The results of the Wald test show that there is no heterogeneity for the pre-reform period but there is heterogeneity for the reform period between the coastal and interior provinces of China. This implies that provincial incomes as a system were converging during the pre-reform period but are diverging during the reform period. The implications of this finding for China's stability and continued reform need no further emphasis.

4. Finally, the insignificance of the investment coefficients suggests that productivity may have played a more important role in explaining the differences of provincial incomes. This result is consistent with a recent IMF report which examines the reasons behind China's rapid growth (see Hu and Khan (2000).

# References


[1] Arellano, Manuel; Bond, Stephen. April 1991,"Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*; 58(2), pages 277-97.

[2] Arellano,Manuel; Bover,Olympia, July 1995,"Another Look at the Instrumental Variable Estimation of Error-Components Models", *Journal of Econometrics*; 68(1), pages 29-51.

[3] Barro, Robert J; Sala-I-Martin, Xavier. 1991, "Convergence across States and Regions", *Brookings Papers on Economic Activity*; 0(1), pages 107-58.

[4] Barro, Robert-J.; Sala-i-Martin, Xavier. April 1992, "Convergence", *Journal of Political Economy*; 100(2), pages 223-51.

[5] Barro, Robert-J; Sala-i-Martin,-Xavier. 1995, "Economic growth", Advanced Series in Economics. New York; London and Montreal: McGraw-Hill, pages xviii, 539

[6] Baumol, William-J, December. 1986, "Productivity Growth, Convergence, and Welfare: What the Long-run Data Show", *American Economic Review*; 76(5), pages 1072-85.

[7] Baumol,William-J.; Wolff, Edward-N. December 1988, "Productivity Growth, Convergence, and Welfare: Reply", *American Economic Review*; 78(5), pages 1155-59.

[8] Ben David, Dan and Michael B. Loewy, 1997, "Free Trade, Growth and Convergence", NBER Working Paper No. 6095.

[9] Bernard, A. B. and Durlauf, S. N. 1996, "Interpreting Tests of the Convergence Hypothesis" *Journal of Econometrics* 71, pages 161-73.

[10] Blundell, Richard; Bond, Stephen. November 1998, "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics*; 87(1), pages 115-43.

[11] Caselli, Francesco; Esquivel, Gerardo; Lefort, Fernando. September 1996, "Reopening the Convergence Debate: A New Look at Cross Country Growth Empirics", *Journal of Economic Growth*; 1(3), pages 363-89.

[12] China Statistics Yearbook, various issues

[13] Chen, Jian; Fleisher,Belton-M. April 1996. "Regional Income Inequality and Economic Growth in China", *Journal of Comparative Economics*; 22(2), pages 141-64.

[14] Cho, Dongchul. Nov.1996, "An Alternative Interpretation of Conditional Convergence Results", *Journal of Money, Credit, and Banking*; 28(4), Part 1, pages 669-81.

[15] Durlauf Steven and Danny T Quah. 1998, "The New Empirics of Economic Growth", Working Paper 6422, National Bureau of Economic Research.

[16] Hsiao,Cheng. 1986,"Analysis of panel data", Econometric Society Monographs, no. 11, Cambridge; New York and Sydney: Cambridge University Press.

[17] Hsueh,Tien tung; Li,Qiang. 1999, "China's National Income", Boulder and Oxford: Westview Press..

[18] Islam, Nazrul. November 1995, "Growth Empirics: A Panel Data Approach", *Quarterly Journal of Economics*;110(4), pages 1127-70.

[19] Hu, Zuliu; Khan, Mohsin. 2000, "Why is China Growing So Fast?", Economic Issues, 8, International Monetary Fund, Washington, D.C.

[20] Jian, Tinalun; Sachs, Jeffrey-D.; Warner, Andrew-M. January 1996, "Trends in Regional Inequality in China", National Bureau of Economic Research Working Paper: 5412.

[21] Judson, Ruth-A.; Owen,Ann-L. January 1997,"Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists", Board of Governors of the Federal Reserve System, Finance and Economics Discussion Paper Series: 1997/03.

[22] Kiviet, Jan-F. July 1995, "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models", *Journal of Econometrics*; 68 (1), pages 53-78.

[23] Lee, Kevin; Pesaran, M. Hashem; Smith, Ron. July-Aug. 1997, "Growth and Convergence in a Multi-country Empirical Stochastic Solow Model", *Journal of Applied Econometrics*;12 (4), pages 357-92.

[24] Lee, Kevin; Pesaran, M. Hashem; Smith, Ron. February 1998, "Growth Empirics: A Panel Data Approach–A Comment", *Quarterly Journal of Economics*; 113(1), pages 319-23.

[25] Mankiw, N. Gregory. 1995, "The Growth of Nations", *Brookings Papers on Economic Activity*; (1), pages 275-310.

[26] Mankiw, N. Gregory; Romer, David; Weil, David-N. "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*; 107 (2), May 1992, pages 407-37.

[27] Naughton, Barry. 1996, "China's Emergence and Prospects as a Trading Nation", *Brookings Papers on Economic Activity*; 0(2), pages 273-337.

[28] Nerlove, Marc. 1999. "Properties of Alternative Estimators of Dynamic Panel Models: An Empirical Analysis of Cross-Country Data for the Study of Economic Growth", in Hsiao, C, Lahiri, K. Lee, L.F. and Pesaran, M.H. *Analysis of Panels and Limited Dependent Variable Models*, Cambridge University Press, pages 136-170.

[29] Nickell Stephen. 1981, "Biases in Dynamic Models with Fixed Effects", *Econometrica*, vol 49, pages 1417-26.

[30] Quah, Danny. 1993," Empirical Cross-Section Dynamics in Economic Growth", *European Economic Review*, 37 (2/3), pages 426-34.

[31] Raiser, Martin. February 1998,"Subsidising Inequality: Economic Reforms, Fiscal Transfers and Convergence across Chinese Provinces", *Journal of Development Studies*; 34(3), pages 1-26.

[32] Ravallion and Jalan (1996), Growth divergence due to spatial externalities. *Economics Letters 53,* pages 227-232

[33] Solow, R.M. 1956, "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Econometrics*, 70, pages 65-94.

[34] Sala-i-Martin, Xavier. November 1994, "Regional Cohesion: Evidence and Theories of Regional Growth and Convergence", Centre for Economic Policy Research, Discussion Paper 1075.

[35] Sala-i-Martin, Xavier-X. July 1996, "The Classical Approach to Convergence Analysis", *Economic Journal*; 106 (437), pages 1019-36.

[36] Sala-i-Martin, Xavier-X. June 1996, "Regional Cohesion: Evidence and Theories of Regional Growth and Convergence", *European Economic Review*; 40 (6), pages 1325-52.

[37] Sevestre, Patrick; Trognon, Alain. 1996, "Dynamic Linear Models", Matyas, Laszlo; Sevestre, Patrick, eds. The econometrics of panel data: A handbook of the theory with applications. Second edition. Advanced Studies in Theoretical and Applied Econometrics, vol 33. Dordrecht; Boston and London: Kluwer Academic, pages 120-44.

[38] Summers, Robert and Alan Heston. 1991, "The Penn World Table ( Mark 5.0): An Expanded Set of International Comparisons, 1950-1988", *Quarterly Journal of Economics*, 106(2), pages 327-68.

[39] Temple, Jonathan. March 1999, "The New Growth Evidence", *Journal of Economic Literature*; 37 (1), pages 112-56.

Table 1: China's Provinces and Geographic Location

| Province | Location | Province | Location |
|---|---|---|---|
| Beijing | Interior | Henan | Interior |
| Tianjing | Coast | Hubei | Interior |
| Hebei | Coast | Hunan | Interior |
| Shanxi | Interior | Guangdong | Coast |
| Inner Mongolia | Interior | Guangxi | Interior |
| Liaoning | Coast | Hainan | Interior |
| Jilin | Interior | Sichuan | Interior |
| Heilongjiang | Interior | Guizhou | Interior |
| Shanghai | Coast | Yunnan | Interior |
| Jiangsu | Coast | Tibet | Interior |
| Zhejiang | Coast | Shaanxi | Interior |
| Anhui | Interior | Gansu | Interior |
| Fujian | Coast | Qinghai | Interior |
| Jiangxi | Interior | Ningxia | Interior |
| Shandong | Coast | Xinjiang | Interior |

Table 2: Average Annual Growth Rates of Per Capita GDP of China's Provinces

| Province | Pre-reform | Reform | % Change |
|---|---|---|---|
| | Period 1952-78 | Period 1978-97 | |
| Beijing | 8.57 | 7.91 | -7.70 |
| Tianjing | 6.07 | 8.53 | 40.53 |
| Hebei | 4.63 | 9.47 | 104.54 |
| Shanxi | 4.73 | 8.16 | 72.52 |
| Inner Mongolia | 3.24 | 8.57 | 164.51 |
| Liaoning | 7.4 | 7.88 | 6.49 |
| Jilin | 3.42 | 8.93 | 161.11 |
| Heilongjiang | 3.73 | 6.70 | 79.62 |
| Shanghai | 7.09 | 8.35 | 17.77 |
| Jiangsu | 2.97 | 12.33 | 315.15 |
| Zhejiang | 3.7 | 13.27 | 258.65 |
| Anhui | 2.46 | 9.22 | 274.80 |
| Fujian | 4.18 | 12.64 | 202.39 |
| Jiangxi | 2.47 | 9.36 | 278.95 |
| Shandong | 5.04 | 10.83 | 114.88 |
| Henan | 3.22 | 9.70 | 201.24 |
| Hubei | 4.41 | 9.64 | 118.59 |
| Hunan | 3.85 | 8.21 | 113.25 |
| Guangdong | 3.94 | 11.51 | 192.13 |
| Guangxi | 4.58 | 8.23 | 79.69 |
| Hainan | N/A | N/A | |
| Sichuan | 3.9 | 9.68 | 148.20 |
| Guizhou | 3.37 | 8.22 | 143.92 |
| Yunnan | 3.83 | 8.99 | 134.73 |
| Tibet | N/A | N/A | |
| Shaanxi | 5.74 | 8.37 | 45.82 |
| Gansu | 3.97 | 7.62 | 91.94 |
| Qinghai | 6.03 | 5.88 | -2.49 |
| Ningxia | 6.77 | 7.21 | 6.50 |
| Xinjiang | 2.68 | 8.99 | 235.45 |
| Average (standard error) | 4.50 (1.567) | 9.09 (1.722) | |

Table 3: Simulation Results for Different Dynamic Panel Data Estimatiors

| N | $b$ | Within | DIFF-GMM | SYS-GMM | OLS |
|---|---|---|---|---|---|
| 100 | 0.5 | -0.0370 | 0.4641 | 0.5100 | 0.8736 |
| | | (0.0697) | (0.2674) | (0.1330) | (0.0266) |
| 100 | 0.8 | 0.1343 | 0.4844 | 0.8101 | 0.9798 |
| | | (0.0726) | (0.8224) | (0.1618) | (0.0107) |
| 100 | 0.9 | 0.1906 | 0.2264 | 0.9405 | 0.9950 |
| | | (0.0725) | (0.8264) | (0.1564) | (0.0055) |
| 500 | 0.5 | -0.0360 | 0.4887 | 0.5021 | 0.8748 |
| | | (0.0310) | (0.1172) | (0.0632) | (0.0114) |
| 500 | 0.8 | 0.1364 | 0.7386 | 0.7939 | 0.9801 |
| | | (0.0328) | (0.3085) | (0.0779) | (0.0047) |
| 500 | 0.9 | 0.1930 | 0.5978 | 0.9043 | 0.9951 |
| | | (0.0330) | (0.6407) | (0.0999) | (0.0025) |

Notes: the table reports means (standard deviations) from experiments with T=4 and 1000 replications. The model is $y_{i,t} = a_i + by_{i,t-1} + u_{i,t}$ with $var(a_i) = var(u_{i,t}) = 1$ and initial conditions drawn from the stationary distribution for $y_{i,1.}$ GMM is two-step estimators.
Source: Blundell and Bond (1998).

Table 4: Cross-sectional Tests for Unconditional Convergence

| | Dependent Variable: $\ln y_1 - \ln y_0$ | | | |
|---|---|---|---|---|
| | Pre-reform Period 1953-77 | | Reform Period 1978-97 | |
| a | 0.757 | 0.996 | 2.571* | 3.717* |
| | (0.706) | (0.685) | (-0.808) | (-0.585) |
| $\ln y_0$ | -0.006 | -0.055 | -0.141 | -0.326* |
| | (0.113) | (0.111) | (0.114) | (0.084) |
| Coast | - | 0.197 | - | 0.503 |
| | | (0.104) | | (0.090) |
| $\bar{R}^2$ | 0.001 | 0.056 | 0.055 | 0.548 |

Note: We have 28 observations for all regressions. Standard errors are in parenthesis. * denotes statistically significant at the 5% level. $\ln y_0$ is log of the initial GDP per capita in 1953 or 1978. $\ln y_1$ is log of GDP per capita in 1977 or 1997

Table 5: Cross-sectional Tests for Conditional Convergence

| | Dependent Variable: $\ln y_1 - \ln y_0$ | | | |
|---|---|---|---|---|
| | Pre-reform Period 1953-77 | | Reform Period 1978-97 | |
| Constant | -0.616 | -2.544 | -1.160 | 4.900 |
| | (2.033) | (2.323) | (4.402) | (3.212) |
| $\ln y_0$ | 0.134 | -0.040 | -0.093 | -0.283* |
| | (0.137) | (0.147) | (0.128) | (0.094) |
| $\ln(I/Y)$ | 0.294* | 0.271* | -0.218 | -0.326* |
| | (0.122) | | (0.362) | (0.084) |
| Coast | - | 0.197 | - | 0.503 |
| | | (0.104) | | (0.158) |
| $\ln(n+g+\delta)$ | -0.390 | | -1.266 | |
| | (0.629) | | (1.863) | |
| $\bar{R}^2$ | 0.0956 | 0.293 | 0.0081 | 0.538 |

Notes: see notes for Table 3

Table 6: Panel Data Tests for Conditional Convergence of the Pre-reform Period 1953-77

| | (1) OLS | (2) WG | (3) DIFF-GMM | (4) SYS-GMM |
|---|---|---|---|---|
| ln $(y_{t-1})$ | 1.0248 | 0.3797 | 1.0457 | 0.9795 |
| | (0.0305) | (0.1277) | (0.1213) | (0.0543) |
| ln $(n+g+\delta)$ | -0.1483 | -0.0283 | -0.3558 | 0.0666 |
| | (0.0780) | (0.0771) | (0.1343) | (0.1298) |
| ln $(I/GDP)$ | 0.0134 | -0.1027 | 0.3242 | -0.1724 |
| | (0.0281) | (0.0550) | (0.0956) | (0.0621) |
| *Constant* | -0.3220 | -0.0380 | N/A | 0.2570 |
| | (0.2904) | (0.0346) | (N/A) | (0.3927) |
| *Implied* $\lambda$ | N/A | 19.37% | N/A | 0.414% |
| Sargan *Test* $(p\text{-}value)$ | | | 0.033 | 0.416 |

Note: Standard errors in brackets.

Table 7: Panel Data Tests for Conditional Convergence of the Reform Period 1978-97

| | (1) OLS | (2) WG | (3) DIFF-GMM | (4) SYS-GMM |
|---|---|---|---|---|
| ln $(y_{t-1})$ | 0.9927 | 0.9043 | 1.0299 | 0.8944 |
| | (0.0272) | (0.1765) | (0.0264) | (0.0522) |
| ln $(n+g+\delta)$ | -0.4024 | -0.5676 | -1.1250 | -0.1657 |
| | (0.2451) | (0.2941) | (0.2075) | (0.5964) |
| ln $(I/GDP)$ | -0.0060 | 0.0866 | 0.1411 | 0.1279 |
| | (0.0591) | (0.0770) | (0.0758) | (0.1404) |
| *Constant* | -0.5224 | -0.0422 | N/A | 0.9322 |
| | (0.6031) | (0.0718) | | (1.6535) |
| *Implied* $\lambda$ | 0.15% | 2.01% | 1.11% | 2.23% |
| Sargan *Test* (*p−value*) | | | 0.769 | 0.258 |

Note: See notes for Table 6.

Table 8: Panel Data Tests for Conditional Convergence with Different Technology Progress Rates Between the Coastal and Interior Provinces

| | Pre-reform Period 1953-77 SYS-GMM | Reform Period 1978-97 SYS-GMM |
|---|---|---|
| $\ln(y_{t-1})$ | 0.9570 | 0.8467 |
| | (0.0496) | (0.0303) |
| $\ln(n+g+\delta)$ | 0.0727 | -0.2906 |
| | (0.1325) | (0.4809) |
| $\ln(I/GDP)$ | -0.1735 | 0.0123 |
| | (0.0769) | (0.1183) |
| *Implied* $\lambda$ | 0.879% | 3.328% |
| *Wald test* | 0.303 | 0.000 |
| Sargan *Test* | 0.541 | 0.163 |

*Note: Standard errors in brackets. The figure reported for the Wald test is the p-value of the null hypothesis of joint insignificance of the coast dummy interacted with time dummy. The figure reported for the Sargan test is the p-value of the null hypothesis of valid instruments.*

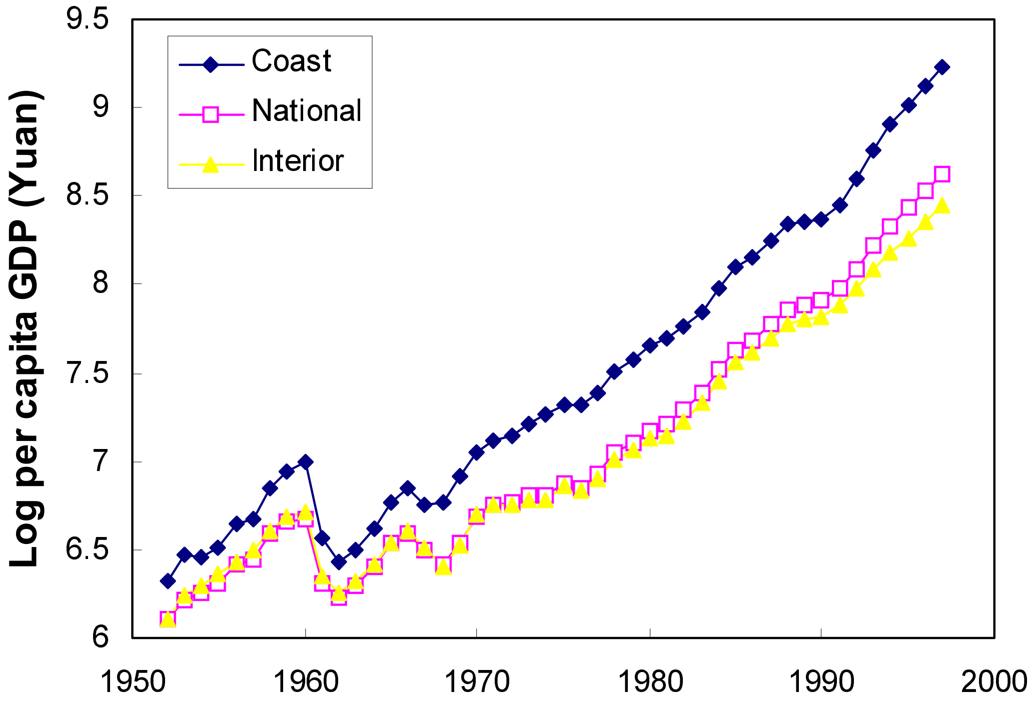


Figure 1: Log GDP per capita of National, Coastal Provinces and Interior Provinces 1952-97

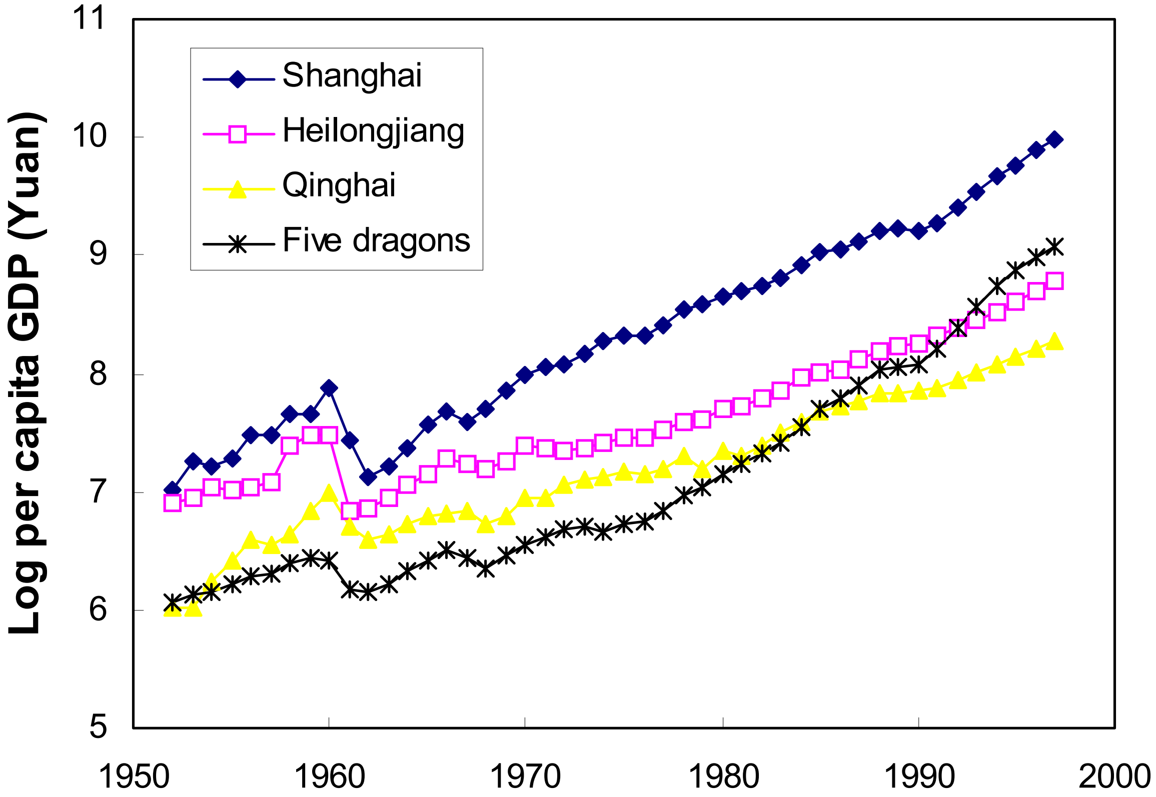


Figure 2: Five Dragons, Heilongjiang and Qinghai Comparing with Shanghai in terms of Per Capita GDP

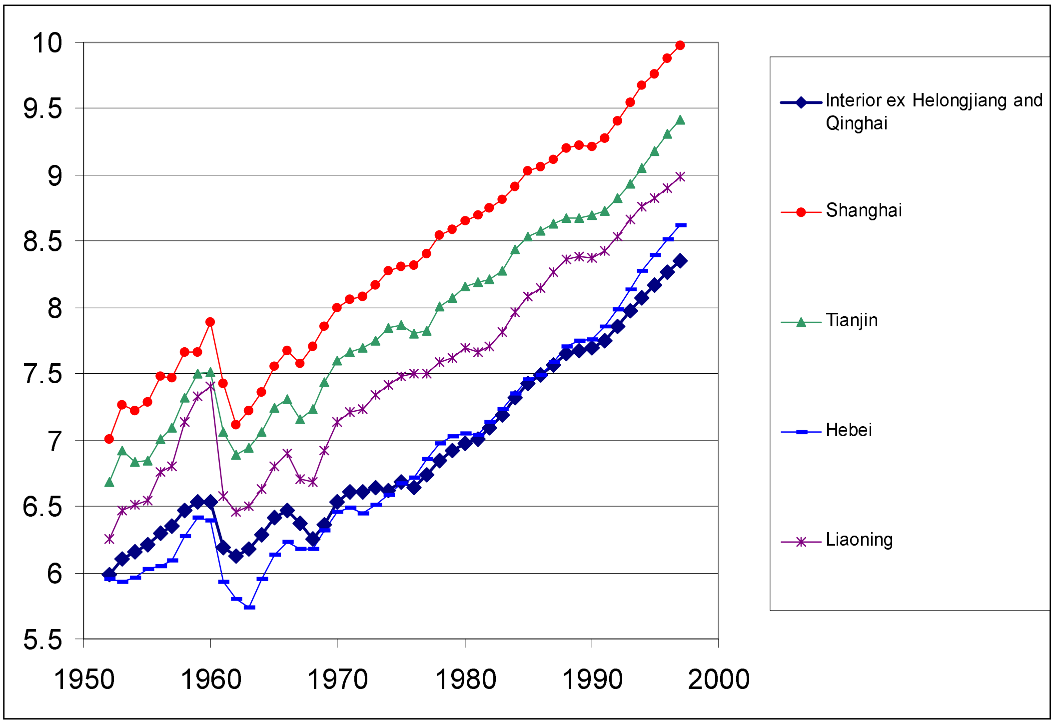


Figure 3: Log GDP per capita of Four Coastal Provinces (Hebei, Liaoning, Tianjing and Shanghai) and 17 Interior Provinces Excluding Heilongjiang and Qinghai

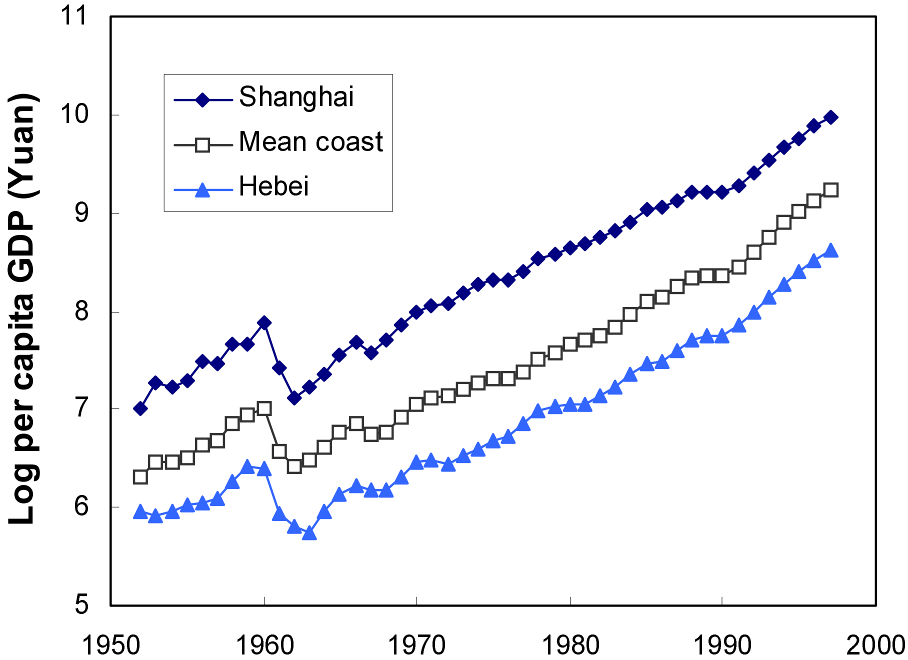


Figure 4: Log GDP per capita of the Richest (Beijing), the Poorest (Guizhou) and Mean of Interior Provinces of China 1952-97

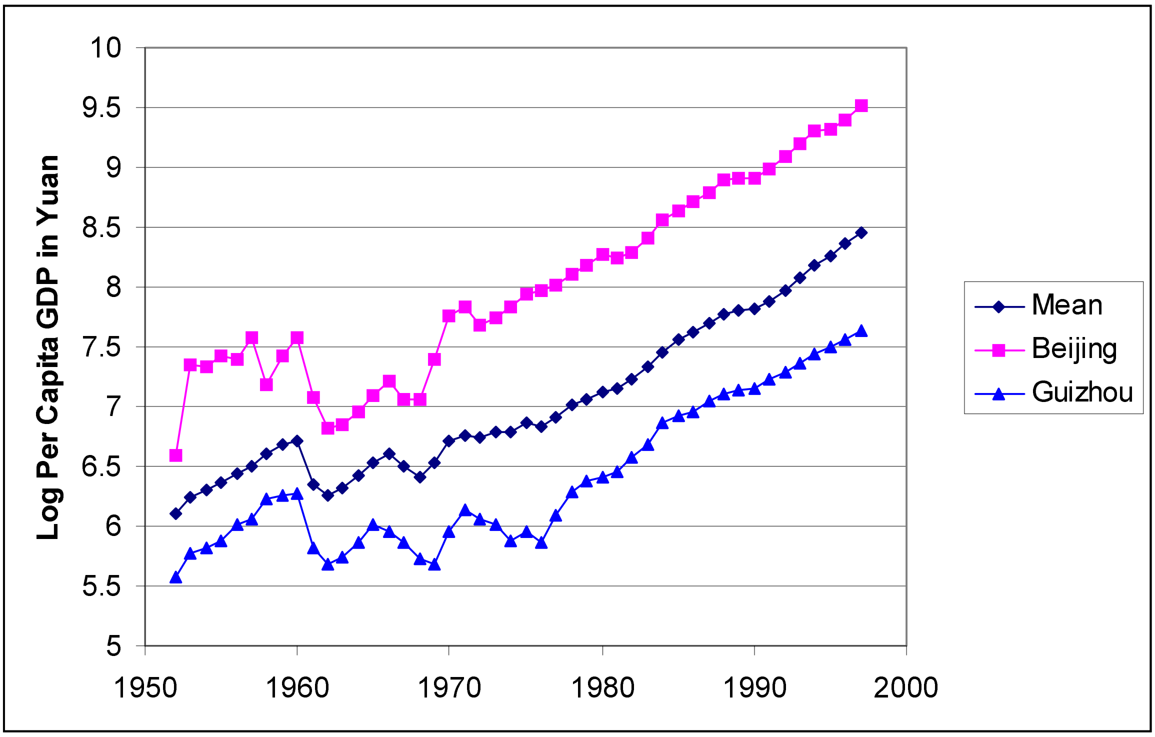


Figure 5: Log GDP per capita of the Richest (Beijing), the Poorest (Guizhou) and Mean of Interior Provinces of China 1952-97