Are China's Post-reform Provincial Income Levels Diverging?[¤]

Peter Pedroni[y] — Indiana University & Williams College

James Yudong Yao — International Monetary Fund

June 10, 2002

Abstract

Recent policy studies have argued that conditions have prevailed in China since the open door economic reforms the late 1970's that have encouraged rapid growth at the expense of regional income inequality across the provinces of China. However, many empirical studies based on provincial level data have concluded that although per capita incomes appeared to have been diverging prior to the reforms, they have actually been converging since the time of the reforms. In this paper we argue that these conclusions need to be reconsidered in light of recent improvements in both the data and the empirical methodologies that are now available. By applying tests that explicitly account for the time series properties of the data, we …nd that the recently released Hsueh-Li (1999) data supports the conclusion that the open door reforms set o¤ a period in which per capita incomes have been diverging among the Chinese provinces. Furthermore, we show empirically that this national divergence among provinces cannot be attributed simply to the presence of separate regional convergence clubs among the coastal versus interior provinces.

[¤] Acknowledgements: We thank Yin-Wong Cheung and Menzie Chinn for helpful suggestions.

[y] Corresponding author. Please address correspondence to Peter Pedroni, Department of Economics, Indiana University, Bloomington, IN 47401, ppedroni@indiana.edu . After August 1, 2002, please address correspondence to Peter Pedroni, Department of Economics, Williams College, Williamstown, MA 01267.

1 Introduction

China's experience with economic growth has changed dramatically following the reforms initiated in 1978 which introduced economic incentives and opened the economy to foreign trade and investment. In the years prior to the reforms, real per capita income grew at 2.5% per annum. Growth during this period was also quite erratic, with dramatic variations associated with the "Great Leap Forward" and the breaking of trade relations with the Soviet Union in 1958-60. The second half of the 1960's witnessed agricultural failures and a nation-wide famine caused by the political and social chaos that followed the Cultural Revolution. By contrast, the reforms since 1978 have unleashed a period of unprecedented rapid and steady growth, which has 8.8% per annum from 1978-1997 for the economy as a whole, and 10.2% for the coastal provinces excluding the city provinces. The contrast in the growth experience between these two periods can be seen readily in Figure 1, which depicts real log per capita income for the average of the 28 provincial level localities for which data is available going back to 1952.

In general terms, the pre-reform period 1952-77 is characterized by unsteady growth, with relatively large fluctuations in measured per capita income. These are due in large part to two distinct historical episodes. The first is the "Great Leap Forward" campaign from 1958 to 1960 and the subsequent agricultural failures and a nation-wide famine. The second is a large decline in the latter half of the 1960s due to the political and social chaos that followed the initiation of the Cultural Revolution from 1965 to 1968. The early 1970s was then characterized by a recovery period, and by the late 1970s the growth trend of the Chinese economy improved dramatically and has remained high since. The takeoff toward higher growth coincides with the process of economic reform and open door policy adopted in 1978. However, associated with the rapid growth, there has also been a dramatic increase in personal income inequalities stemming from the increased emphasis on market incentives, and the reversal of the "iron rice bowl" policies. On the basis of official statistics, the World Bank (1997) reports an increase in the Gini coefficient from 28.2 in 1981 to 38.8 in 1995. Inequalities in rural areas have been growing fastest although urban inequality has also risen sharply in the most recent period between 1988 and 1995.[1]

It should not be surprising that the transition from socialism to more market-oriented policies has increased income inequality at the household level. But the fact that systematic income inequalities also appear to be increasing at the inter-provincial level is somewhat more perplexing. The issue is of central concern to the Chinese authorities, as it bears directly on the success of decentralization policies and the political cohesion of the country at the national level. In September 1995 the Chinese government endorsed the view that regional inequalities have widened since the reforms stating that "since the adoption of reforms and open door policies, we have encouraged some regions to develop faster and get richer, advocated that the richer should act as a model for and help the poor. Each region has had immense economic development and the people's standard of living has had great improvement. But for some reason, regional economic inequalities have widened somewhat."[2]

Recently, Young (2000) has also provided compelling arguments and empirical evidence to support the idea that despite the fact that China has liberalized international trade, interprovincial trade has actually become more restrictive in the post reform period, and that this combination can be expected to generate large regional disparities. However, previous empirical studies have for the most part come to the opposite conclusion. For example, Chen and Fleisher (1996), Jian, Sachs and Warner (1996) and Raiser (1998) generally …nd the absence of per capita income convergence among provinces during the period from 1952-1977, followed by a pattern of convergence during the post reform period. Jian, Sachs and Warner (1996) argue that convergence among the provinces of China has been a relatively recent phenomenon, emerging strongly only since the reform period began in 1978.[3]

In this paper, we argue that empirical results on Chinese provincial convergence should be reevalu-

[1] See for example Ravallion and Jian (1999), Kahn, Gri¤en and Risken (1999) and Yang (1999) for recent discussion of these trends in income inequality.

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

[3] Only …fteen provinces in their study have GDP data for the pre-reform period 1952-77.

ated in light of more recently available data and in light of recent advances in empirical methodologies for testing convergence. Specifically, we make use of the newly released national income data set of Hseuh-Li (1999), and apply empirical techniques that explicitly account for the time series properties of the data that emphasize the long term trending behavior of the data. These techniques allow us to obtain a fairly dramatic picture of the trends toward and away from convergence in the two periods prior to and following the economic reforms. In contrast to previous studies that have relied on conventional cross section techniques, we find evidence that accords with the more recent arguments put forth by Young, as well as the pragmatic observations of policy makers in China. Specifically, in this study we find strong evidence to support the idea that although real per capita incomes were generally converging among provinces prior to the economic reforms begun in 1978, the reforms initiated a period in which provincial real per capita incomes have since been diverging rather than converging. Furthermore, we also investigate the relative growth patterns among regional subgroups, as well as among subgroups of provinces that have received differing degrees of preferential open door policy support. On the basis of this analysis, we show that the national divergence cannot easily be explained simply on the basis of separate regional or political convergence clubs.

The remainder of the paper is organized as follows. In section 2 we describe the data set on provincial Chinese incomes and the measurement issues involved in constructing an appropriate measure for our study. In section 3 we take a brief informal look at the raw data to understand the subtleties involved in determining whether per capita incomes are indeed converging or diverging among Chinese provinces in the pre and post reform periods. Next in section 4 we set up a more formal criteria for convergence of provincial real per capita incomes and describe the empirical methodology that we employ to formally test and compare income convergence properties before and after the reforms based on recently developed techniques for nonstationary panel data analysis. Section 5 discusses and evaluates the results that we obtain, and section 6 provides some concluding remarks.

2 Chinese Provincial Income Data

The population of China has exceeded 1.2 billion since the 1990s and covers an area of 9.6 million square kilometers. Administratively, The Peoples Republic of China was founded in 1949 and is divided into 30 provincial-level localities in the mainland, including 22 provinces, three municipalities directly under the central government, and …ve autonomous regions[4]. Unless otherwise stated, we will hereafter refer to all of these provincial-level localities simply as provinces. The average population size of these provinces is 40 million, which is substantially larger than most states and provinces of other large countries. Table (1) provides the list and geographic location of China's provinces.

By and large, the Chinese statistical reporting system is set up under a well organized hierarchical framework, from the top national level, known as the State Statistical Bureau (SSB), down to provincial, county, and township levels. Essentially, these form a regular reporting system for the most local units to submit their statistical data to the statistical authorities at the national level.

During the pre-reform period of 1952-77, the Chinese statistical system adopted the system of material product balances (MPS). The design of the MPS accounting system was originally tailored to meet the needs of the central planning economy. The main aggregate indicators are total output value of society and national income. National income is value-added and comparable with GDP in the system of national accounts (SNA) adopted by the market-economy countries. The main di¤erence between GDP and national income is that national income does not account for the value-added of the service sector.

Since the adoption of the comprehensive reform policy in 1978, China's statistical system has had to adapt to the wave of reforms. Starting from 1985, the State Statistical Bureau in China has received …nancial aid from the World Bank and the Asian Development Bank to further develop the national income accounts system, particularly with respect to the estimation of GDP and its components in the

[4] The two special administration regions, Hong Kong and Macao and the island of Taiwan are not included in this study.

SNA as well as to correct deficiencies of MPS accounts. The SNA system was then implemented on a nationwide basis in 1992. While the provincial statistical data of the pre-reform period was largely absent until recently, it is now becoming increasingly available.

Chinese provincial level data is normally compiled from both an original source and a secondary source. The original sources derive from three different Chinese official statistics publications. They are (A) the China Statistics Yearbook from 1981 to 1997, with issue 1982 absent, (B) the China Historical Statistics 1949-1989 and (C) various volumes of the 28 provincial statistics yearbooks. Sources A and B are national sources and offer the advantage of methodological consistency across provinces. Moreover, source B is aimed at building up historical data by province, so that data for each province contain the same variable and are calculated in the same fashion. Therefore, source B plays a key role in all provincial statistics. Source C provides more detailed data over a longer time, but often differs from province to province in terms of years listed, or other factors. So direct comparison across provinces is not as viable on the basis of source C. The main secondary source has been the Hsueh-Li-Liu (1993) data set. This data set provides a comprehensive set of provincial data drawn from source B and C, covering the period 1949-89. The authors have attempted to use a consistent methodology in compiling these data, thereby overcoming some of the potential comparability problems However, there are three key shortcomings of these data sets which have plagued previous empirical work. First, there is no provincial GDP data for about half of China's provinces during the pre-reform period. Instead, most previous studies on China's regional economy have used NI data at the province level. Second, there are no provincial GDP deflators. Therefore, the provincial retail price index is typically used to deflate NI for each province.

Recently, with the support and cooperation of the SSB, Hsueh and Li (1999) have made significant progress in improving the Chinese statistical data. They have published the most complete set of Chinese national income from 1952 to 1995 based on the SNA both at the national level and provincial level. Two contributions of their recent work are particularly valuable for provincial level analysis. First, by combining historical statistical data that had been stored at the SSB and existing provincial national income data, they were able to add in estimates for the service sector that had been omitted in the MPS accounts and thereby construct provincial GDP data for the pre-reform period 1952-77 which are more directly comparable with those of the reform period of 1978-95. Second, they provide estimates of real provincial GDP, investment growth rates and the corresponding de‡ators for the period 1952-95.

Based on these newly published estimates, the data that we use consists of provincial GDP for 28 provinces of China for a 45 year period from 1952 to1997. Speci…cally, we use the provincial GDP data from Hsuech and Li (1999) for the period 1952-95 and use provincial GDP data for 1996 and 1997 obtained from the 1997 China Statistics Yearbook. In order to study income convergence, we use total population in combination of GDP and de‡ated by provincial GDP de‡ators to generate real per capita GDP by province. The GDP de‡ator takes the prices of 1995 as 100%. There are two geographic areas that are excluded from the data. One is Hainan, which is a newly established province, and the other is Tibet for which data is missing prior to 1987.

3 Informal Analysis

The most important outcome of China's economic reform and open-door policy is its remarkable economic growth performance. During the reform period China has been growing at a spectacular rate, enabling per capita income to more than quadruple. Today China is among the fastest growing economies of the world. Table 2 shows that the average annual growth rate of real per capita GDP of all provinces during the reform period 1978-97 is 9.09%.

The acceleration of economic growth of China's provinces has been very broadly based. With the exception of Beijing and Qinghai, all provinces' average annual rates of the reform period are higher than those of the pre-reform period. In particular, there are seven provinces, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Henan and Xinjiang, which have more than doubled their average annual growth rates during the post-reform period relative to the pre-reform period. However, average annual growth rates over the last two decades of the reform period vary significantly across the provinces and as reported in table 1, the standard deviation of the provincial growth rates during the reform period is higher than during the pre-reform period. The slowest growing province, Qinghai, only grew at an annual average of 5.88% during the post-reform period, though needless to say, this still represents a respectable growth rate relative to most economies or regions of the world. Heilongjiang, which is one of the old Soviet style heavy industry based provinces has been the second slowest in terms of growth during the post-reform period, at 6.7%. Qinghai's and Heilongjiang's growth performances stand in a sharp contrast to the five fastest growing coastal provinces, which have managed to grow at double digit rates[5]. We refer to these five provinces as the "five dragons". However, in geographical terms, it is not the case that all of the coastal provinces have been growing faster than the interior provinces. For example, the remaining four coastal provinces, Tianjin, Hebei, Liaoning and Shanghai, annually grow at 8.53%, 9.47%, 7.88% and 8.35%, respectively, during the post-reform period, and among these, only Hebei exceeds the national average provincial growth rate of 9.09%. There are five interior provinces, Anhui, Jiangxi, Henan, Hubei and Sichuan, which outperform the three coastal provinces Tianjin, Liaoning and Shanghai. Hence a coastal location alone appears not to be a sufficient condition for the highest growth rates.

Next, we illustrate by graphical analysis some of the key features of the provincial growth process during the pre-reform and reform periods based on the Hsueh and Li data. These illustrations help us to gain a broad sense of the general tendencies toward income convergence or divergence among China's 28 provinces during both the pre-reform period 1952-77 and the reform period 1978-97 before implementing more formal tests for convergence.

[5] These provinces are Jiangsu, Zhejiang, Fujian, Shandong and Guangdong, with the average annual growth rates over 1978-97, respectively, given by 12.33%, 13.27%, 12.64%, 10.83% and 11.51%.

One way to characterize the provincial growth pattern is to take Shanghai, the richest, as the national technology leader or frontier, and identify provinces that are catching up to Shanghai, those that are falling behind Shanghai, and those which are roughly growing together with Shanghai. Among the coastal provinces, the group known as the …ve dragons, Guangdong, Fujian, Jiangsu, Shandong and Zhejiang, have grown at a rate of over 10% each year since 1978. Figure 2 illustrates the extend to which the average per capita income of these …ve dragons is rapidly approaching that of Shanghai. The …gure also shows that two provinces have most clearly been lagging behind the rest. These two provinces are Heilongjiang, the Soviet-style heavy industrial base in the northeast of China, and Qinghai, an inland province in the northwest of China. By the late 1980's these two provinces were overtaken by the …ve dragons provinces. Figure 3 depicts how, in contrast to these eight provinces, the remaining interior provinces appear to be growing along distinct but roughly parallel growth paths.

Many studies of the Chinese economy tend to categorize provinces simply into coastal versus interior groups, since historically there has been a signi…cant di¤erence in the growth rates between these. However, the Hsueh and Li data show that this dual classi…cation may be insu¢cient in that it masks considerable individual heterogeneity within the categories. For example, …gure 4 shows that the log per capita income gap between the coastal and interior provinces increased only slightly after 1978. By contrast, when we compare the wealthiest and poorest of each of these classi…cations with the group averages of the remaining provinces we see a di¤erent picture. For example in …gure 5 we see that the income gap between the group average and the richest coastal province, Shanghai, and the poorest coastal province, Hebei, is relatively constant over time. This appears to contradict the idea that the coastal provinces behave as a single convergence club. Similarly, …gure 6 shows that the income gap between the richest interior province, Beijing, and the poorest interior province, Guizhou, has even increased slightly since the 1980s.

If we think of convergence as the reduction of interprovincial income di¤erences, it appears that there is relatively little overall convergence among the group of all provinces when examined over the entirety of the sample period. Furthermore, even as possible regional growth clubs, the coastal provinces and interior provinces fail to converge within each group. As each of the …gures illustrates, per capita growth rates, re‡ected in the slopes of the logged series, while di¤ering across provinces, appear relatively stable over time since the1978 reforms. Accordingly, the di¤erences in growth rates appear to be fairly persistent, which implies that interregional income disparities are likely to continue to diverge.

The graphical analysis is useful in developing informally a sense of the extent of convergence or divergence present in the data. To reconcile whether the data is actually consistent with long run convergence or divergence in a formal sense, we next turn to a more systematic empirical analysis.

4 Panel Based Tests for Long Run Convergence

To investigate whether the data is consistent with long run convergence or divergence more precisely, we begin with a formal de…nition of what we intend by the concept of long run convergence in panels such as the Hsueh and Li data. In particular, we employ a de…nition of income convergence in keeping with the one studied in Evans (1998) for an international panel of country level data. This notion of convergence asks whether or not the long-run forecasts for output di¤erences converge as the forecasting horizon increases, which implies that the long run income gap between any two provinces must be stationary.

An important implicit distinction between this empirical formalization and the informal graphical analysis of the previous section is that here in order to conclude in favor of convergence, we require that the properties of the data must be consistent with the fact that di¤erences are eliminated eventually, and not necessary that the di¤erences are becoming smaller at all points in time. Clearly this less restrictive concept for convergence is somewhat more subtle, and far more di¢ cult to detect on the basis of a casual graphical analysis. To formalize this idea empirically, for provincial income data which individually exhibit nonstationarities, one can characterize this criteria for convergence as follows. Suppose that y_{it}, the logarithm of per capita output for province i at time t, is difference stationary, and thus exhibits unit root behavior individually. Then any pair of provinces i and j are said to converge pairwise if the difference $y_{it} - y_{jt}$ is stationary so that y_{it} and y_{jt} are cointegrated. Convergence between members of a larger group of provinces is then defined analogously by requiring that every pair within the set exhibits convergence. Note furthermore that if the stationary differences between provinces have nonzero means, then this corresponds to the notion of conditional convergence, since the convergence is said to be conditional upon the province specific fixed effects. One obvious advantage to focusing in this way on the properties of long run income gaps as the criteria for convergence is that is directly allows us to infer long run forecasts for the absence or presence of income inequalities between provinces.

For the group of Chinese provinces, one might imagine testing this condition pairwise for all provinces within the sample and then requiring that the condition hold for each possible pair of provinces. An obvious disadvantage of such an approach is that conventional tests for cointegration tend to have low power for such short samples, and so the probability of failing to reject the null of no cointegration for at least some pairs would be quite high regardless of the true relationship. Fortunately, as Evans (1998) demonstrates, it is possible to translate this criteria into a single criteria that should apply to the group as a whole when interpreted as a panel. Specifically, Evans shows that the criteria of pairwise convergence for all members of panel is equivalent to the condition that the difference between the individual series, y_{it}, and the mean value for the series across all members at each point in time, $\bar{y}_t = \frac{1}{N}\sum_{i=1}^{N} y_{it}$, is stationary. Thus, the condition states that all members converge pairwise if $y_{it} - \bar{y}_t$ is stationary for each member $i = 1, 2, ..., N$ of the panel. Consequently, Evans argues that the null of nonconvergence can be interpreted as the unit root null in panel unit root test.

Consequently, in this context, whether or not convergence is occurring can be evaluated by asking whether or not the autoregressive parameter β_i is zero for the panel data regression given by

$$\Delta(y_{it} - \bar{y}_t) = \mu_i + \beta_i(y_{i,t-1} - \bar{y}_{t-1}) + \sum_{k=1}^{K_i} \phi_{i,k} \Delta(y_{i,t-k} - \bar{y}_{t-k}) + \varepsilon_{it} \qquad (1)$$

for $i = 1, 2, ..., N$; $t = 1, 2, ..., T$. Notice that this specification is essentially an augmented Dickey-Fuller regression applied to the panel of income differentials between the individual provinces and the mean income value of the provinces as a group. In this case the μ_i fixed effects represent the individual province's average sample difference from the group mean $(y_{it} - \bar{y}_t)$, which is permitted to vary by province. The autoregressive parameter for the income differentials, β_i, becomes the key coefficient for determining the presence or absence of convergence, the lagged difference terms are intended to capture higher order serial correlation in the time series process for income differentials and the number of lags, K_i, are chosen in a manner to ensure that the remaining error terms ε_{it} are serially uncorrelated. Under this specification, rejection of the panel unit root null hypothesis

$$H_0 : \beta_i = 0 \text{ for all } i$$

in favor of the alternative hypothesis

$$H_1 : \beta_i < 0 \text{ for some } i$$

implies that at least some subset of the members of the panel are converging toward one another. By contrast, failure to reject this null can be taken to imply that no subset of the members of the panel are converging toward one another.

To test this hypothesis, we employ the panel unit root tests of Im, Pesaran and Shin (1997), Maddala and Wu (1999). These tests have an advantage over earlier generation tests such as Breitung and

Meyer (1994), Quah (1994) and Levin and Lin (1993) in that they allow for greater ‡exibility under the alternative hypothesis. Speci…cally, both the Im, Pesaran and Shin (1997) and Maddala and Wu (1999) tests allow the value for the autoregressive coe¢cient, β_i, under the alternative hypothesis to vary across provinces. By contrast, the Levin and Lin (1993) panel unit root test employed in the …rst generation of panel unit root convergence tests such as Evans (1998) and Evans and Karras (1996) require the autoregressive coe¢cient to be homogeneous under the alternative hypothesis, so that $\beta_i = \beta < 0$. Thus, the more recent tests provide us with the additional ‡exibility of allowing the convergence dynamics to di¤er across provinces under the alternative hypothesis, which is clearly an advantage in the current context.

The distinction is achieved by the di¤erence in the way the data are pooled to construct the statistics. For example, while the earlier tests such as Levin and Lin (1993) were based on pooling along the "within" dimension of the data, the Im, Pesaran and Shin (1997) tests are constructed by pooling along the "between" dimension of the data. In practice this means that the Im, Pesaran and Shin tests are constructed on the basis of averaging the unit root tests for the individuals to produce a group mean test statistic. In one such test, Im, Pesaran and Shin recommend constructing a t-bar statistic, which is based on averaging the individual augmented Dickey-Fuller unit root t-tests. Consequently, to compute the statistic in our context, one …rst estimates the augmented Dickey-Fuller regression given in equation (1) above individually for each of the $i = 1, 2, \ldots, 28$ provinces of the panel and then constructs the 28 corresponding ADF t-statistics, t_i. These individual statistics are averaged to obtain the t-bar statistic $\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_i$. Finally, since the distribution for the individual ADF t-statistics are not centered around zero under the unit root null hypothesis, it becomes necessary to adjust for this feature to ensure that the distribution of the t-bar statistic does not diverge under the null hypothesis as the number of individual members of the panel, N, grows large. Fortunately, under the null hypothesis, the mean of the individual t_i is a known constant as the sample size T grows large, as is the standard deviation of the individual t_i. Consequently, the t-bar statistic is adjusted by subtracting off the mean and dividing by the standard deviation, so that the statistic becomes

$$\tilde{z} = \sqrt{N}(\bar{t} - u)/s$$

where u is the known mean of the individual ADF t-statistic distribution, and s is the known standard deviation of the individual ADF t-statistic distribution. Provided that the individual statistics are independent, then as Im, Pesaran and Shin demonstrate, this statistic will be distributed as standard normal under the null hypothesis, and will diverge to negative infinity under the alternative hypothesis so that large negative values can be taken to reject the null hypothesis.

Maddala and Wu (1999) suggest a somewhat different approach to testing the unit root null hypothesis in panels, which also allows the more flexible modeling approach of the Im, Pesaran and Shin test under the alternative hypothesis as compared to the earlier panel unit root tests. Specifically, Maddala and Wu suggest that rather than basing the pooled test statistic on the average value of the individual member test statistics, one can also base the panel unit root test on pooled values of the marginal significance level associated with the individual member test statistics. Since the marginal significance levels, or "p-values" p_i, for the individual tests are uniformly distributed between 0 and 1, this implies that $-2 \log p_i$ is distributed as a χ^2 with two degrees of freedom. For the N members of the panel, we can sum these values to obtain the Pearson-lambda statistic, also commonly referred to as the Fisher statistic, which becomes

$$P_\lambda = -2 \sum_{i=1}^{N} \log p_i$$

Again, under the assumption that the individual statistics are independent, the sum of N independent χ^2 with two degrees of freedom implies that the Fisher statistic is distributed as a χ^2 with $2N$ degrees of freedom. When a sufficient number of the individual p-values are small enough to indicate rejection of the null hypothesis, then the Fisher statistic takes on increasingly large values. Thus, large positive values in the right hand tail of the $\chi^2_{(2N)}$ distribution indicate rejection of the unit root null hypothesis.

The cost of the Maddala and Wu approach is that the distributions for the individual ADF based unit root tests are nonstandard and depend on Brownian motion functionals, which means that the p-values for arbitrary values of the t-statistics are not available in tabular form, and must be simulated. Consequently, the approach can be very computer intensive, particularly if one wishes to condition the p-values on the sample size and on the particular lag truncation that is fitted. On the positive side, Maddala and Wu argue that their approach does well in pointing out the role of individual members in contributing to the overall results for the panel. In particular, it is useful to know whether the results for the panel are generally being driven by the strength of one or two outlier members or whether it is a general tendency of all members of the panel. For example, it is quite likely that in practice that panels of the type that we are considering here are mixed, in the sense that within either the pre-reform or post-reform periods, it is possible that not all members behave as if they are stationary or nonstationary. Instead, the panels may be mixed in the sense that the majority of the members contain unit roots but one or two do not. In this case, we would like to know the basis for the rejection of the panel as a whole. By computing the marginal significance levels of the members individually prior to pooling the results, we can get a sense of whether or not this is the case.

Finally, the marginal significance analysis associated with the Maddala and Wu approach brings with it another important benefit. Both the Im, Pesaran and Shin t-bar test and the Maddala and Wu Pearson-lambda test are constructed under the assumption that the income differentials are independent of one another across provinces. But in practice, it is possible that even these differentials contain feedback effects that render them dependent upon one another across provinces. In this case, these tests are not strictly valid. Fortunately, once the marginal significance levels have been computed, it is also possible to use these to produce a more conservative test that is also invariant to the presence of cross sectional dependency. As Maddala and Wu point out, the marginal signi…cance levels can be used to construct a test based on the Dufour and Torres (1996) criteria for the Bonferroni inequality constraint, which does not require independence across the individual members of the panel. The Bonferroni inequality constraint indicates that the marginal signi…cance level P for a rejection of the null hypothesis H_0 applied to the panel of N members is given by $P \quad \sum_{i=1}^{N} p_i$ where, as before, the p_i are the marginal signi…cance levels for the tests applied to the individual members. Dufour and Torres recommend using the criteria that $p_i = \frac{P}{N}$, to set the rejection level. To understand the nature of the test, consider the following. Imagine that we are interested to know whether we should reject the null hypothesis for the panel at the 10% level. Clearly, it would be a mistake to conclude that we should reject the hypothesis simply because a single member of a panel of N provinces produces a p-value less than 10%. To do so would ignore the fact that for a panel of N members, we would expect to reject at the 10% level $\frac{N}{10}$ times regardless of whether the null is false. Thus, the Dufour and Torres criteria for the Bonferroni inequality corrects for this by indicating that conservatively we should only reject the null hypothesis for the panel if any one individual indicates a rejection level of $\frac{P}{N}$ or stronger. The strength of the rejection in this form is not altered by whether or not the individual test results are correlated and thus the test is invariant to cross sectional dependency across provinces.

Taken together, the Im, Pesaran and Shin (1997) and Maddala and Wu (1999) approaches provide a nice opportunity to investigate both the individual and group provincial dynamics in terms of the convergence hypothesis and to evaluate the contribution of individual provinces to the results for the panel as a whole. Consequently, in what follows we combine the two approaches to form a uni…ed analysis of the convergence properties of per capita income in the Chinese provinces.

5 Empirical Results

In this section we present and discuss the results of the formal convergence tests described in the previous section. We divide the sample of 28 provinces into a pre-reform subsample, from 1952 through 1977, and a post-reform subsample, from 1978 through 1997. We also consider various province subgroupings to investigate the possibility of convergence clubs in the post-reform period.

To begin, the IPS t-bar test and Maddala-Wu Fisher test are applied to equation (1) for the pre-reform period 1952-77 and the post-reform period 1978-97, respectively. Table ?? summarizes the results for the pre-reform period and post-reform periods. In keeping with the discussion of the previous section, the lag truncations for the individual ADF unit root regressions were allowed to vary by individual province in both subsamples, for both the individual tests as well as the panel based test. In each case, the lag length was chosen by a standard data dependent step down procedure, which is typically implemented for the ADF unit root test in conventional time series regressions. Speci…cally, the step down procedure involves starting with a su¢ ciently large number of lags and then sequentially eliminating the highest order lags one at a time until one of them tests signi…cant. In our case, we allowed this step down procedure to choose a di¤erent lag truncation for each province. For the arbitrary initial starting value, we rounded o¤ to the nearest integer of 1/5 of the sample length. Thus, for the pre-reform period, with T=26, we started with an initial "maximum" lag value of 5, and then allowed the automated data dependent procedure to choose the actual number of …tted lags, which then varied between 0 and 5. For the post-reform period, with T=20, we started with an initial "maximum" lag of 4 so that the actual number of …tted lags then varied between 0 and 4. Since both individual and panel unit root tests are well known to be sensitive to the number of lags …t, we also experimented with using maximum lag truncations that varied from 6 to 2 in the case of the pre-reform period and 5 to 2 in the pre-reform period. The results for the panel were not altered by these choices. Consequently, in the interest of space, we report only the tables with results for the case with maximum truncations of 5 and

4 respectively. For the IPS tests, we conditioned the individual mean and variance adjustment terms on both the sample size and the lag truncation value that was chosen endogenously for each individual.

Consider …rst the results reported in table (??) for the pre-reform period. The …rst column to the right of the province name reports the value for the individual ADF t-statistic for the particular province. The next column reports the associated marginal signi…cance level, also known as the "p-value", for the reported ADF t-statistic. As discussed in the previous section, since the distribution for the ADF t-statistic is nonstandard, the p-values must be estimated by Monte Carlo simulation. Furthermore, since the size of the test under the null hypothesis is very sensitive to the sample size, as well as the number of …tted lags, we used a bootstrap to condition the Monte Carlo simulation on both the sample size and the speci…c number of lags that were …tted in each case. In all cases, the simulation was based on 20,000 draws from a pure random walk of length T+100. The …rst 100 realizations of each random walk were discarded to reduce the impact of arbitrary initial conditions, and the ADF regression was then …tted with the number of lags that had been …tted according to the data dependent step down procedure.

Notice that based on these p-values, we see that only three provinces were able to reject the unit root null at the 5% level or better, and only 6 were able to reject at the 10% level or better. On the other hand, we can also see from the p-values that evidence was not necessarily an even split. In other words, although not many provinces provided small enough p-values to support rejections on their own at the 5% or 10% level, a great many provinces were able to support rejections at say the 35% level or better. Individually, these would not be taken as su¢cient evidence. However, the combined evidence of these marginally supportive provinces is su¢cient to produce a rejection for the panel as a whole. This is precisely what the Fisher statistic of Maddala-Wu test does. It reports the combined evidence, and supports a rejection of the unit root null hypothesis at a value of 79.35 since under the unit root null hypothesis it should be distributed as a Â^2 with 56 degrees of freedom, which implies that the test value of 79.35 corresponds to a p-value of 0.022 as reported in the table. Likewise, the Im, Pesaran and Shin T-bar statistic produces a value of -2.23, which supports a rejection of the unit root null hypothesis, since it is distributed as standard normal under the unit root null hypothesis, which implies that the test value of -2.23 corresponds to a p-value of 0.013 as reported in the table. The use of the individual p-values helps us to confirm that this rejection is not based on the unusually strong results of only one or two outliers. Consequently, these tests support the conclusion that inter-provincial per capita incomes were converging in the sense that incomes were converging toward one another for a significant subset of provinces during the pre-reform period.

Next, consider the results for the post-reform period, which are reported in table (??). For this sub-sample, we encounter a very different situation. Although two provinces are able to reject at the 5% level or better, and four are able to reject at the 10% level or better, the pattern for the majority of provinces is now much different. Rather than being close to rejections, the majority of the test statistics are nowhere near the left tail of the distribution, and the p-values reflect this. Now, even the combined evidence does not reject the null hypothesis, and both the Fisher statistic and the Im, Pesaran and Shin t-bar statistic reflect this. In both cases, the statistics are far from rejecting the null hypothesis. This points to the likelihood that on balance, the majority of the provinces are not converging to one another in the post-reform period since we cannot reject the panel unit root null hypothesis for the differences $(y_{it} - \bar{y}_t)$. On the other hand, the fact that a small subset of provinces do provide rejections leads us to consider the possibility that there may be subgroups for which convergence may be present. We also note that at least one province provides a rejection that is marginally consistent with the Bonferroni test at the 5% level, since $p_i = \frac{\alpha}{N} = \frac{0.05}{28} = 0.0018.$

Consequently, we next consider that the possibility that the apparent absence of convergence in the post-reform period nationally can be attributed to the idea that at least some subsets of countries are converging to separate regional or policy determined clubs. Notice that the results for the full sample of provinces already indicate that this result is unlikely, since they indicate that we cannot reject the likelihood that there is no sizeable subset of provinces which converge pairwise within the sample. Nevertheless, given the presence of sampling variation and the fact that we cannot say a priori what constitutes a sufficiently sizeable subset on the basis of the full sample results, it is worth further investigating the convergence properties of candidate subsets of provinces. For example, many researchers have proposed that differences in geography or differences in preferential open door policies at the provincial level may generate convergence clubs among the provinces of China. The most commonly proposed of these has been the coastal versus interior geographic distinction. Recently, Demurger et al (2001) have also investigated by more traditional methods whether differences in the level of preferential economic policy at the provincial level may help to explain patterns of growth among the provinces. We consider both of these possibilities.

Toward this end, we first examined various regional subgroupings for the possibility of geographically based convergence clubs. The designation for coastal versus interior tends to vary among studies, and so we have experimented with a number of different coastal versus interior classifications as well as other regional subgroupings. As described in table 2, strictly speaking, our sample consists of 10 provinces that lie along the coast of China. Among these, the southern autonomous coastal region of Guangxi is sometimes excluded from the coastal designation and grouped with the other southwestern interior provinces of Sichuan, Yunnan and Guizhou. Likewise, the northeastern coastal province of Liaoning is occasionally excluded from the coastal designation and grouped along with the other interior Manchurian provinces of Heilongjiang and Jilin, each of which had heavily industrialized Soviet style economies during the pre-reform period and have tended to lag behind in the post-reform period. Finally, Tianjin and Shanghai are also sometimes excluded from the coastal group since they represent somewhat unique metropolitan areas. We experimented with each of these different coastal classifications. In the interest of space, we report in table form only two classifications, the broadest classification including all 10 provinces, and the narrowest classification, including only the 6 non-metropolitan central coastal provinces. In table 4 we report both the individual and group results for each of these two benchmark coastal groupings. Notice that, statistically, the individual test values for the same province tends to differ depending on the grouping in which it is included. This is because when we test $(y_{it} - \bar{y}_t)$, the value for $\bar{y}_t$. differs depending on which other provinces are included in the group. In all cases, for each of the coastal groupings, including the other coastal groupings not reported in table form, we were unable to reject the null hypothesis that all of the $(y_{it} - \bar{y}_t)$ are nonstationary on the basis of any of the tests. In other words, the evidence strongly contradicts the presumed presence of a separate coastal convergence club regardless of which combination of coastal provinces we consider.

We also investigated the convergence properties for the interior provinces. The first benchmark is for all of the remaining 18 interior provinces of our sample which do not have a coastline and is reported in the first two columns of results in table 5. The results for the interior provinces as a group is somewhat mixed. The IPS tests fail to reject, while the Fisher test and several individual provinces reject at the 5% level. The Bonferroni test easily rejects at the 1% level. This leads us to suspect that while the interior provinces as a whole do not represent a convergence club, if there is a convergence club, it is likely to lie among some subset of these interior provinces. Thus, we next subdivided the interior provinces into various smaller geographic subsets, such as the 6 central interior provinces, the central interior plus northeastern provinces, the central interior plus southwestern interior provinces, the 6 northwestern interior provinces, and the interior provinces minus the northwestern interior provinces. These results proved to be interesting. All of the groupings which excluded the northwestern interior provinces uniformly failed to reject the null. In these cases, the IPS, Fisher and Bonferroni statistics were always in agreement. By contrast, the grouping that included only the interior northwestern provinces was the only one of the groupings for which rejections of the null were obtained in some cases. Specifically, for the group of northwestern provinces, the IPS test rejected at the 10% level with a p-value of 0.074, while the Fisher statistic came close with a p-value of 0.123, although the Bonferroni did not reject. Collectively, these results show that if a case is to made for a geographically based convergence club among the Chinese provinces during the post-reform period, at best the only possible candidate appears to be the more geographically isolated provinces of the northwestern interior. The individual and group test statistics for two of these interesting benchmark cases are reported in the remaining columns of table 5, namely the subgrouping of the 6 northwestern provinces and the subgrouping of the interior provinces with the northwestern ones excluded.

Next, we investigated various province subgroupings based on the extent to which preferential open door policies have been extended in the post-reform period. For this, we used the index constructed in Demurger et al (2001). Specifically, Demurger et al construct an index ranging from 0 to 3 for each province during the post-reform years depending on the type and extent of favored free trade zones that are present, and report the average index value for each province over the post-reform period. These average values range from 0.33 to 2.86, with the majority of provinces below 1.0. We divided these into three roughly equal quantiles. This produced 9 "low preference" provinces with average index values below 0.5, plus 9 "medium preference" provinces with values between 0.5 and 1.0, and 10 "high preference" provinces with values exceeding 1.0. The assignment of provinces into these groups is also described in table 1. As it turns out, the high preference quantile coincides exactly with the subset of all coastal provinces. Therefore, we already know based on the geographically defined classifications that this subset does not contain a converge club, as reported in table 4. As expected, the result continues to hold even when we exclude the most extreme high preference provinces, Fujian and Guangdong, which are the only two provinces with average index values in excess of 2.0.

Similarly, we already know from this that the results for the remaining 18 medium and low preference provinces are mixed, since they coincide with our benchmark group of all interior provinces. In effect, therefore, the preferential policy classification system primarily provides us with an alternative way in which to further decompose these 18 interior provinces, along policy groupings rather than geographical groupings. When we do this by examining the low preference and medium preference quantiles separately, we find another interesting result, as reported in table 6. In this case, the low preference provinces continue to produce weak or mixed results primarily in the direction of a nonrejection. For example, while the Fisher and Bonferroni reject, the IPS clearly does not reject. Among the individual province tests in this group, there are clearly two outliers: Gansu, which lies at the extreme left tail of the distribution with a p-value close to zero, and Qinghai, which lies at the extreme right tail of the distribution with a p-value close to one. When either or both of these northwestern provinces are excluded, the panel results uniformly fail to reject the null, leading us to conclude that it is unlikely that a significant convergence group lies within this subset. By contrast, it is the medium preference provinces which provide us with the most likely candidate for a convergence club. For this group of provinces, all of the tests are in agreement in rejecting the null hypothesis. The IPS, Fisher and Bonferroni tests all reject at the 5% level or better. These are by far the strongest results in favor of a rejection of the null among the many subgroups that were considered. Consequently, if we are to look for a policy based converge club in the post-reform period, it is not among the most open high preference coastal provinces, nor is it among the least favored of the interior provinces. Rather, it appears to be among the middle of the road provinces, which are neither wide open to international trade and investment nor which are relatively shut off from international trade and investment.

Taken together, the results paint an interesting picture for the growth pattern among the Chinese provinces in the post-reform period. As a general phenomenon, per capita incomes among the provinces do not appear to be converging, but rather appear to be diverging from one another in the post-reform period. Furthermore, it does not appear to be the case that this can be explained by the presence of a simple dual convergence club that distinguishes between coastal and interior provinces. On the contrary, per capita incomes in the coastal provinces do not appear to be converging toward one another regardless of which coastal provinces we consider. Likewise, the interior provinces as a group also do not appear to be converging to each other, but are in general diverging. At most, there is some indication that the more geographically isolated subset of these provinces, consisting of the 6 northwestern interior provinces, may not be diverging from one another, so that these may represent a small subset of interior provinces that are on a common convergence path. Similarly, when we examine the growth patterns among provinces with similar degrees of preferential open door policies, we …nd that the quantile with the most preferential treatment are growing along divergent paths. The least preferentially treated quantile also does not appear to provide much evidence for convergence as a group. Rather, the middle quantile, with a moderate ranking for preferential policies appears to be the most likely to contain a convergent subgroup. This quantile varies geographically among the interior provinces, and includes several of the northwestern provinces along with a majority of central provinces.

Broadly speaking, many of these patterns might arguably be viewed as being roughly consistent with the interpretations expressed in Young (2000) and elsewhere, namely that the opening up of international trade, while the degree of interprovincial trade and factor mobility remains relatively limited, has allowed individual provinces to grow along divergent paths. Young (2000) concludes that " There is every indication that the economy of the People's Republic, while opening up internationally, has become fragmented internally." Under this interpretation, relatively limited interprovincial trade in the presence of accelerated international trade has restricted the degree of interprovincial convergence, and in most cases has lead to diverging per capita incomes across provinces. In terms of the growth patterns among the subgroupings of provinces, we …nd that those which been permitted to open up most to international trade have been able to grow fastest, and have also grown along their own separate paths. Those which remain least open to trade have grown least, and with little internal interprovincial trade to induce convergence, have remained on separate nonconvergent growth paths. If there is any pattern of convergence among the provinces, it is a middle quantile phenomenon. In particular, it is the subset of provinces which have neither grown exceptionally from the opening up of international trade nor stagnated from the lack of it that may be following along a similar convergence path in the post-reform period. Consequently, the evidence appears to support the idea that further extending preferential open door policies to other provinces may lead to further acceleration of the growth process, but it is more likely the freeing up of internal factor movements that will facilitate convergence and a lessening of interprovincial disparities in per capita incomes.

6 Concluding Remarks

China's accelerated growth rate during the reform period 1978-97 has reinforced concerns about how to cope with continued growth while also maintaining balanced regional income inequality. This paper examines the issue of Chinese provincial income convergence issue using the new provincial income data set from Hsueh and Li (1999), which span both the pre-reform and post-reform periods of Chinese economic growth. By using recent econometric developments which account for the time series properties of the data, our primary results indicate that although at least a significant subset of China's provincial incomes appear to have been converging toward one another during the pre-reform period, subsequent to the reforms initiated in 1978, this no longer appears to be the case. Rather, per capita incomes in the majority of provinces appear to be diverging from one another since 1978. In contrast to many earlier empirical studies, our findings support the view of Chinese government and of the UNDP that interprovincial inequalities have been widening since 1978. The findings are also consistent with casual observation, and are consistent with the implications of restrictive interprovincial trade raised in Young (2000) and Demurger et al (2001). Furthermore, the evidence does not provide support for the idea that the provincial patterns of growth can easily be explained on the basis of a geographically oriented dual convergence club consisting of the usual coastal versus interior classification.

Historical experience elsewhere in the world 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 inequalities. As such the existence of regional income disparities is of considerable interest as it bears directly on the sustainability of economic reform and open door policy. The results of this study add to a growing body of literature providing further con…rmation that there could be a risk that growing regional income inequality might derail the stability of China's economic reform process and imperil prospects for future growth if it is not addressed.

References

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

[2] Breitung, Jorg and Wolfgang Meyer, (1994) "Testing for Unit Roots in Panel Data: Are Wages on Di¤erent Bargaining Levels Cointegrated?", Applied Economics, 26, pages 353-361.

[3] China Statistics Yearbook,1997.

[4] Demurger, Sylvie, Je¤rey Sachs, Wing Thye Woo, Shuming Bao, Gene Chang and Andrew Mellinger (2001) "Geography, Economic Policy and Regional Development in China," forthcoming in Asian Economic Papers.

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

[6] Evans, Paul (1998) "Using Panel Data to Evaluate Growth Theories", International Economic Review; 39(2), pages 295-306.

[7] Evans, Paul and Georgios Karras (1996) "Convergence Revisited", Journal of Monetary Economics; 37(2), pages 249-65.

[8] Hsueh,Tien Tung, Qiang Li and Shucheng Liu, eds. (1993) "China's provincial statistics, 1949-1989", Boulder and Oxford: Westview Press, pages xxiii, 595.

[9] Hsueh,Tien Tung and Qiang Li (1999) "China's National Income", Boulder and Oxford: Westview Press, pages xxiii, 609.

[10] Im, Kyung-So, Hashem Pesaran and Yongdul Shin (1997) "Testing for Unit Roots in Heterogeneous Panels", Discussion Paper, University of Cambridge.

[11] Jian, Tinalun, Je¤rey D. Sachs and Andrew M. Warner (1996) "Trends in Regional Inequality in China", National Bureau of Economic Research Working Paper: 5412, pages 24

[12] Levin, Andrew and Chien-Fu Lin (1993) "Unit Root Test in Panel Data: New Results", Working Paper, Economics Department, UCSD, San Diego.

[13] Maddala G. S. and Shaowen Wu (1999) "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", Oxford Bulletin of Economics and Statistics, 61(4), 631-52.

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

[15] World Bank (1997) "China Regional Disparities," World Bank Report No. 14496-CHA.

[16] Young, Alwyn (2000) "The Razor's Edge: Distortions and Incremental Reform in the People's Republic of China"; Quarterly Journal of Economics, 115(4),1091-1135.

References

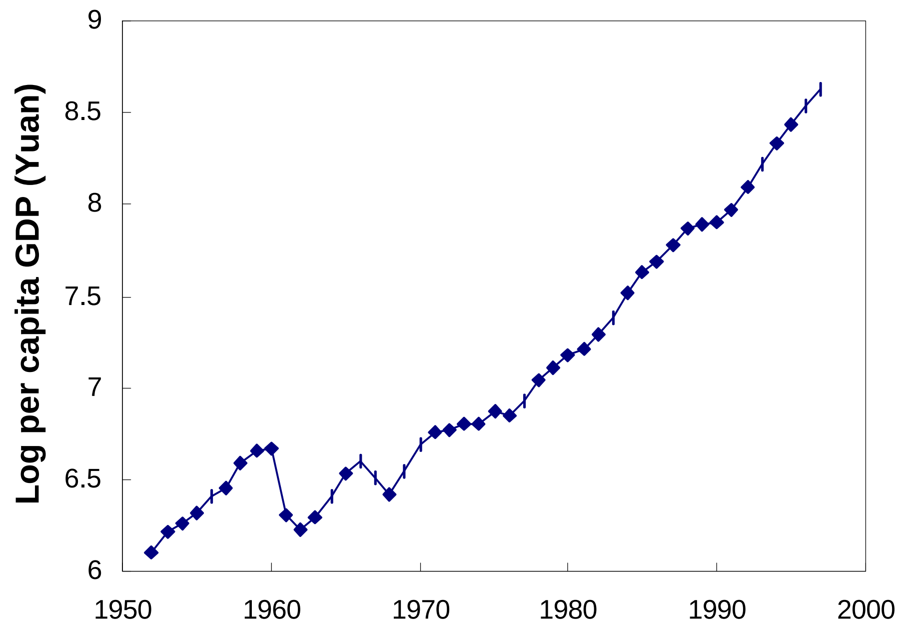

Figure 1: Average provincial log per capita GDP in constant 1995 Yuan prices

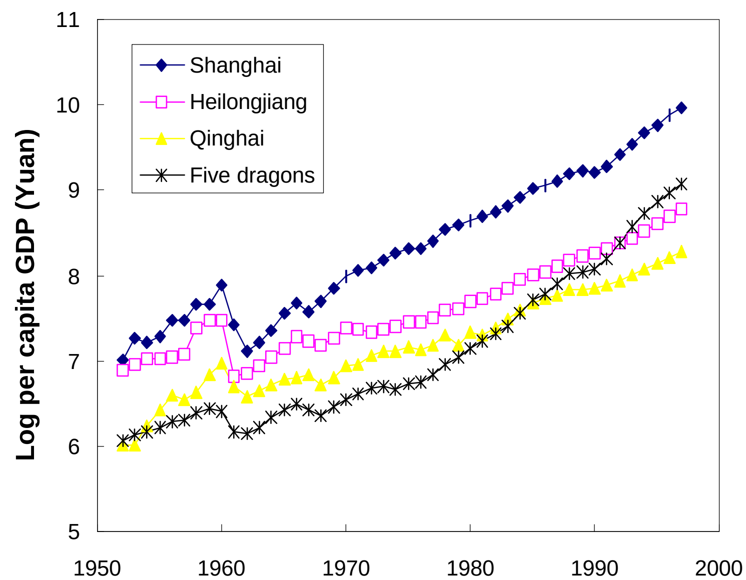

Figure 2: Average per capita incomes of the …ve fastest growing coastal provinces relative to Shangai and slow growing provinces

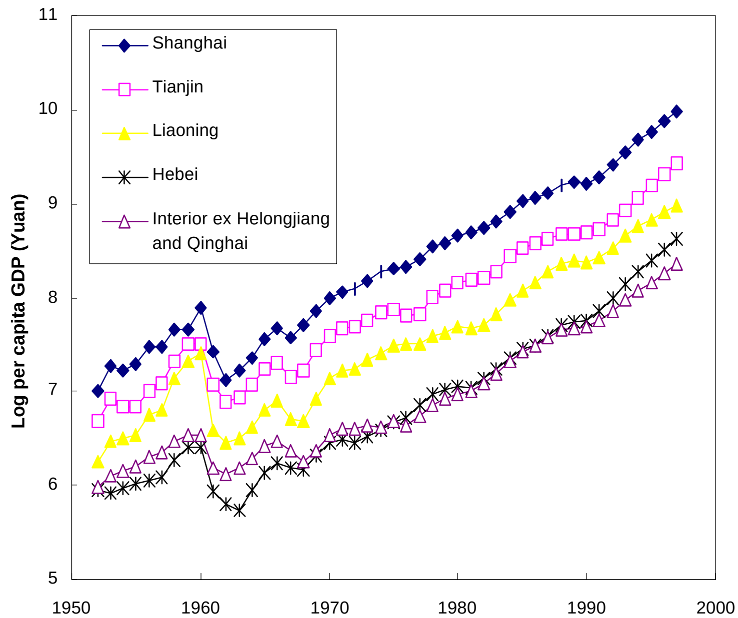

Figure 3: Average interior provincial per capita incomes excluding Qinghai and Heilongjiang relative to Shanghai and others

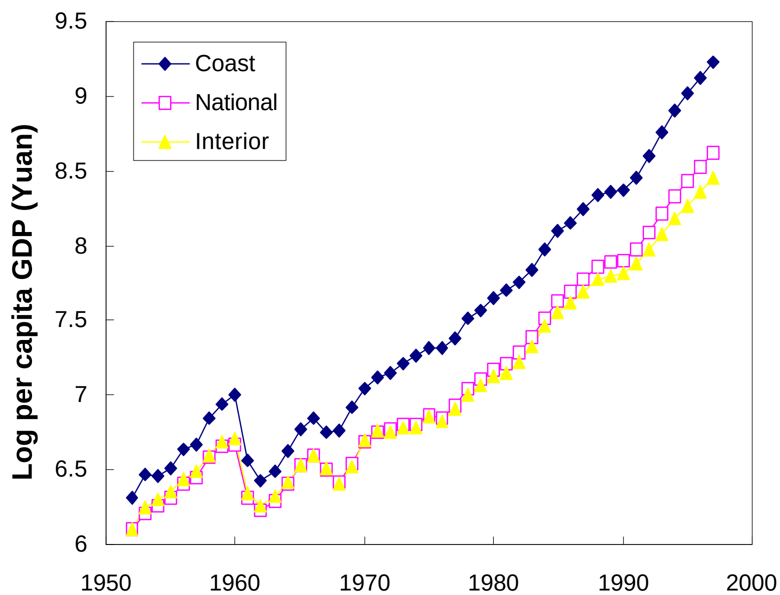

Figure 4: Average coastal versus interior provincial incomes relative to the national mean

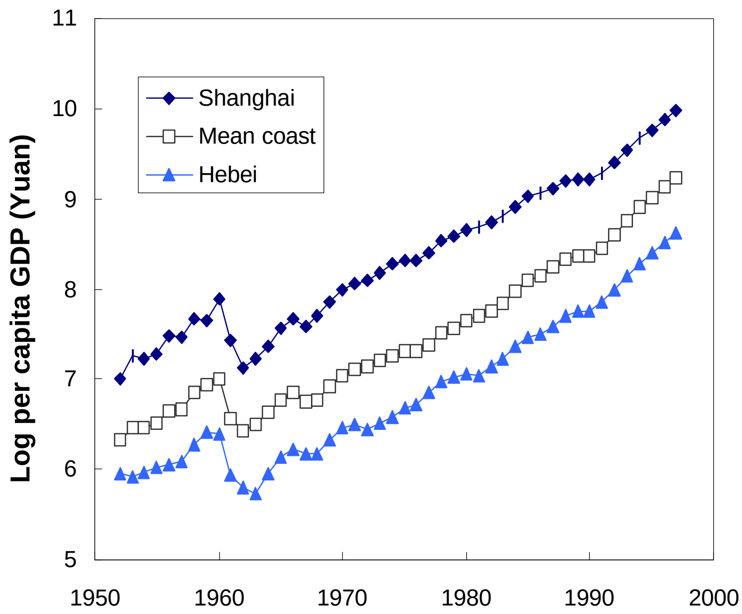

Figure 5: Roughly parallel growth paths of the poorest and wealthiest coastal provinces relative to average

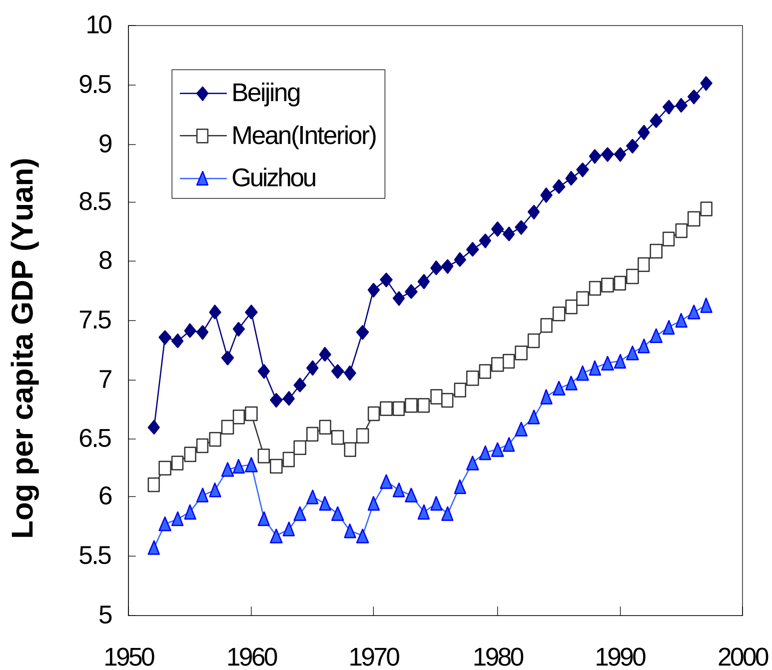

Figure 6: Roughly parellel growth paths among the wealthiest and poorest interior provinces relative to average

Table 1: List of China's Mainland Provinces and Geographic Location

Code	Province	Location	Pref. Level
1	Beijing	Interior-Central	medium
2	Tianjing	Coastal-Central	high
3	Hebei	Coastal-Central	high
4	Shanxi	Interior-Central	low
5	Inner Mongolia	Interior-NW	medium
6	Liaoning	Coastal-NE	high
7	Jilin	Interior-NE	medium
8	Heilongjiang	Interior-NE	medium
9	Shanghai	Coastal-Central	high
10	Jiangsu	Coastal-Central	high
11	Zhejiang	Coastal-Central	high
12	Anhui	Interior-Central	medium
13	Fujian	Coastal-Central	high
14	Jiangxi	Interior-Central	low
15	Shandong	Coastal-Central	high
16	Henan	Interior-Central	low
17	Hubei	Interior-Central	medium
18	Hunan	Interior-Central	low
19	Guangdong	Coastal-Central	high
20	Guangxi	Coastal-SW	high
21	Hainan	NA	high
22	Sichuan	Interior-SW	medium
23	Guizhou	Interior-SW	low
24	Yunnan	Interior-SW	medium
25	Tibet	NA	NA
26	Shaanxi	Interior-NW	low
27	Gansu	Interior-NW	low
28	Qinghai	Interior-NW	low
29	Ningxia	Interior-NW	low
30	Xinjiang	Interior-NW	medium

Table 2: Provincial Growth Performance during Pre-reform and Post-reform Periods

Province	Average Annual Growth Rate of Per Capita GDP	
	Pre-reform Period 1952-77	Reform Period 1978-97
Beijing	8.57	7.91
Tianjin	6.07	8.53
Hebei	4.63	9.47
Shanxi	4.73	8.16
Inner Mongolia	3.24	8.57
Liaoning	7.4	7.88
Jilin	3.42	8.93
Heilongjiang	3.73	6.70
Shanghai	7.09	8.35
Jiangsu	2.97	12.33
Zhejiang	3.7	13.27
Anhui	2.46	9.22
Fujian	4.18	12.64
Jiangxi	2.47	9.36
Shandong	5.04	10.83
Henan	3.22	9.70
Hubei	4.41	9.64
Hunan	3.85	8.21
Guangdong	3.94	11.51
Guangxi	4.58	8.23
Hainan	N/A	N/A
Sichuan	3.9	9.68
Guizhou	3.37	8.22
Yunnan	3.83	8.99
Tibet	N/A	N/A
Shaanxi	5.74	8.37
Gansu	3.97	7.62
Qinghai	6.03	5.88
Ningxia	6.77	7.21
Xinjiang	2.68	8.99
Average (standard deviation)	4.50 (1.567)	9.09 (1.722)

Figure 7: Table 3. Comparison of Pre- and Post-Reform Periods

Province	pre-reform period ADF	pre-reform period p-value	post-reform period ADF	post-reform period p-value
Beijing	-2.56	0.116	-0.55	0.862
Tianjin	-1.49	0.522	-1.58	0.476
Hebei	-1.21	0.650	-0.43	0.788
Shanxi	-1.86	0.343	-0.50	0.873
Inner Mongolia	0.19	0.920	-1.52	0.467
Liaoning	-1.96	0.302	-1.37	0.571
Jilin	-2.35	0.168	-2.37	0.165
Heilongjiang	-2.71	0.088	0.10	0.890
Shanghai	-1.03	0.727	-3.09	0.046
Jiangsu	-2.29	0.185	0.08	0.957
Zhejiang	-2.74	0.083	-0.18	0.889
Anhui	-1.90	0.326	-2.37	0.156
Fujian	-0.18	0.856	1.79	0.997
Jiangxi	-1.70	0.422	-1.95	0.286
Shandong	-0.94	0.760	0.73	0.989
Henan	-2.93	0.050	-2.96	0.061
Hubei	-0.51	0.801	-2.25	0.188
Hunan	-3.49	0.018	-1.17	0.667
Guangdong	-3.42	0.021	-0.85	0.770
Guangxi	-1.88	0.336	-5.12	0.001
Sichuan	-1.94	0.311	1.26	0.982
Guizhou	-0.72	0.820	1.17	0.996
Yuannan	-2.12	0.242	-2.21	0.214
Shaanxi	-2.92	0.058	-3.00	0.053
Gansu	-2.25	0.198	-0.62	0.741
Qinghai	-2.47	0.126	2.90	1.000
Ningxia	-1.21	0.653	-0.20	0.923
Xinjiang	-1.78	0.358	-1.48	0.521
panel IPS-tbar	-2.23	0.013	3.45	1.000
panel MW-Fisher	79.3	0.022	56.9	0.442

Figure 8: Table 4. Post Reform Coastal/High Preference Subgroupings

Province	all coastal group ADF	p-value	nonmetro-central-coast ADF	p-value
Tianjin	-0.35	0.899	NA	NA
Hebei	-2.33	0.158	-1.44	0.509
Liaoning	0.41	0.926	NA	NA
Shanghai	-0.84	0.783	NA	NA
Jiangsu	-0.48	0.864	-2.54	0.125
Zhejiang	-1.71	0.384	-2.32	0.160
Fujian	1.08	0.988	0.51	0.963
Shandong	-0.82	0.792	-1.01	0.731
Guangdong	-1.31	0.600	-1.49	0.517
Guangxi	-1.60	0.469	NA	NA
panel IPS-tbar	2.10	0.982	0.18	0.571
panel MW-Fisher	9.78	0.972	11.20	0.512

Figure 9: Table 5. Post-Reform Interior Geographic Subgroupings

Province	all interior subgroup ADF	p-value	interior minus NW ADF	p-value	Northwest only ADF	p-value
Beijing	-2.40	0.156	-1.75	0.394	NA	NA
Shanxi	-1.41	0.557	-1.11	0.692	NA	NA
Inner Mongolia	-5.29	0.001	NA	NA	-1.39	0.525
Jilin	-1.67	0.432	-1.81	0.368	NA	NA
Heilongjiang	-1.27	0.524	-0.98	0.739	NA	NA
Anhui	-0.82	0.792	-0.97	0.742	NA	NA
Jiangxi	0.06	0.954	-0.12	0.935	NA	NA
Henan	-1.28	0.616	-1.33	0.589	NA	NA
Hubei	-1.06	0.711	-1.36	0.582	NA	NA
Hunan	-4.16	0.007	-1.95	0.304	NA	NA
Sichuan	-0.77	0.759	-1.09	0.596	NA	NA
Guizhou	0.35	0.975	1.05	0.995	NA	NA
Yuannan	-1.36	0.582	-1.45	0.537	NA	NA
Shaanxi	-1.47	0.529	NA	NA	-1.95	0.304
Gansu	-5.24	0.001	NA	NA	-2.51	0.124
Qinghai	2.81	1.000	NA	NA	-3.08	0.046
Ningxia	-1.05	0.716	NA	NA	-1.38	0.574
Xinjiang	-1.80	0.371	NA	NA	-2.06	0.263
panel IPS-tbar	-0.07	0.473	1.58	0.943	-1.45	0.074
panel MW-Fisher	52.70	0.036	12.75	0.970	17.75	0.123

Figure 10: Table 6. Post-Reform Medium and Low Preference Subgroupings

	medium preference		low preference		low pref. minus outliers	
Province	ADF	p-value	ADF	p-value	ADF	p-value
Beijing	-1.68	0.426	NA	NA	NA	NA
Shanxi	NA	NA	-2.03	0.275	-1.56	0.488
Inner Mongolia	-2.66	0.098	NA	NA	NA	NA
Jilin	-2.38	0.162	NA	NA	NA	NA
Heilongjiang	-0.93	0.650	NA	NA	NA	NA
Anhui	-2.07	0.243	NA	NA	NA	NA
Jiangxi	NA	NA	0.01	0.949	0.37	0.976
Henan	NA	NA	-1.07	0.707	-1.12	0.687
Hubei	-1.22	0.640	NA	NA	NA	NA
Hunan	NA	NA	-2.27	0.160	-3.39	0.029
Sichuan	-4.64	0.003	NA	NA	NA	NA
Guizhou	NA	NA	-0.76	0.808	0.41	0.978
Yunnan	-1.83	0.358	NA	NA	NA	NA
Shaanxi	NA	NA	-1.55	0.491	-1.36	0.582
Gansu	NA	NA	-7.07	0.000	NA	NA
Qinghai	NA	NA	2.60	1.000	NA	NA
Ningxia	NA	NA	-1.19	0.657	-1.24	0.633
Xinjiang	-1.54	0.495	NA	NA	NA	NA
panel IPS-tbar	-1.95	0.026	0.23	0.592	1.17	0.879
panel MW-Fisher	29.95	0.038	27.35	0.073	11.37	0.657