Convergence across Chinese Provinces: An Analysis using Markov Transition Matrix

Hiroshi Sakamoto, Research Assistant Professor, ICSEAD
and
Nazrul Islam, Research Professor, ICSEAD

Working Paper Series Vol. 2005-14
October 2005

The International Centre for the Study of East Asian Development, Kitakyushu

Convergence across Chinese Provinces:
An Analysis using Markov Transition Matrix

Hiroshi Sakamoto
Nazrul Islam[1]
International Centre for the Study of East Asian Development (ICSEAD)

Abstract

This paper adopts the distribution approach to study convergence across Chinese provinces. In particular, it uses the Markov transition matrix methodology to capture the dynamics embodied in the data and to produce corresponding ergodic distributions. The results indicate that distribution of per capita income across Chinese provinces has become bi-modal over the period of 1952 to 2003. This conforms to the general impression of rising regional disparity in China and is therefore discouraging from the point of view of convergence. However, examination of the dynamics of the post-reform period separately shows that the ergodic distribution on the basis of this period's dynamics is negatively skewed, with more provinces at the higher end of the distribution. This provides some ground for optimism regarding eventual regional convergence in China.

Current draft: October 2005

[1] The authors would like to thank for their useful comments the participants of the ICSEAD seminar where an earlier version of this paper was presented. They are however not responsible for the errors and shortcomings that may remain. Please send your comments to Nazrul Islam, the corresponding author, at nislam@icsead.or.jp

Convergence across Chinese Provinces:
An Analysis using Markov Transition Matrix

1. Introduction

This paper investigates the issue of convergence across Chinese provinces. China represents about one-fifth of the world's population. The population size of most of the Chinese provinces exceeds that of many countries of the world. Whether poorer provinces of China are catching up with richer ones is therefore a question of enormous welfare significance. Another reason why the issue of convergence across Chinese provinces is of interest is the following. China is a country under transition. She is experiencing a switch from central planning to market mechanism, accompanied by a movement away from public ownership over productive assets to private ownership. The neoclassical paradigm of convergence depends on diminishing returns to capital accumulation and diminishing opportunities of technological diffusion. It does not prescribe the institutional setting under which capital accumulation and/or technological diffusion are to occur. The analysis of the Chinese experience should therefore be able to throw light on the question of whether or not changes in the institutional set up is making any difference with regard to the convergence process, and if yes, of what nature.

Most of the studies of Chinese convergence so far have examined either β-convergence or σ-convergence. Both β and σ provide only summary information about the distribution. For example, β provides information regarding the conditional mean, while σ provides information regarding the dispersion. However, these summary statistics cannot reveal what is happening to the shape of the distribution. Yet, introspection and reading of the results of previous studies suggest and the convergence process in China is much more complicated than can be adequately captured by such summary statistics.

In view of the limitations of β and σ, this paper uses the distribution approach advocated by Quah. It uses the Markov transition matrix methodology to capture the dynamics and to reveal the changes in the entire shape of the distribution. The paper uses the data ranging from 1952 to the latest year of 2003 for which data were available. The paper distinguishes two sub-periods, the first being the pre-reform period of 1952-1978 and the second is the reform period of 1978-2003. The analysis is conducted on the basis of both a five-state and a seven-state discretization. Also, to check into the robustness of the results, the paper conducts the analysis on the basis of both annual and five-year transition dynamics.

The results show that the distribution of per capita income across Chinese provinces over time has attained a bi-modal characterisitic. This has been the result of two opposing tendencies prevailing in the two sub-periods. The dynamics of the pre-reform period have worked to make the distribution to be positively skewed, with more provinces piling at lower values of per capita income. The dynamics of the reform period have however worked in the opposite direction, with more provinces moving to higher income groups. Thus although the combined dynamics of the pre- and post-reform periods have yielded a bi-modal distribution, if the dynamics of the post-reform periods hold, the distribution is likely to become more negatively skewed, with more provinces at the upper end of the distribution. In that sense there are grounds for optimism regarding the Chinese convergence process despite the current bi-modality.

The discussion of the paper is organized as follows. Section 2 sets the background by providing a survey of the literature on convergence across Chinese provinces. Section 3 discusses the methodological issues concerning convergence study. Section 4 presents a brief description of the Markov transition matrix methodology for modeling dynamics of distribution. Section 5 presents the cross-section distribution for the beginning year of 1952, the terminal year of 2003, and several important intervening years, such as 1978, the year in which reforms began. These distributions are obtained from actual data using approximation based on the Normal kernel. The main empirical analysis and results are presented in section 6. Section 7 contains the concluding remarks.

2. Convergence study and China

A good part of recent growth research has focused on the issue of 'convergence.'[2] While most of the studies have studied convergence *across countries*, a good number of them have examined convergence *across regions* within a country (often expressed as 'regional convergence').[3] Recent research on *regional convergence* begins with Sala-i-Martin's (1990) study of convergence across states of the USA.[4] Since then this research has expanded to regions of Europe (Barro and Sala-i-Martin 1991), prefectures of Japan (Barro and Sala-i-Martin 1992b), regions of Canada (Coulombe and Lee 1993), counties of Sweden

---

2 For recent surveys of the convergence literature, see Durlauf and Quah (1999), Temple (1999) and Islam (2003).

3 Some studies have considered convergence across regions spread over several countries, as in Europe.

4 This study by Sala-i-Martin was later included in Barro and Sala-i-Martin (1991 and 1992a). Other studies of convergence across US states include Holtz Eakin (1993).

(Persson 1997), states of India (Cashin and Sahay 1995), provinces of Spain (Sanchez-Robles and Villaverde 2001), and even districts of Bangladesh (Hossain 2000). Barro and Sala-i-Martin (1995 and 2004) provide good accounts of the regional convergence literature and the relevant data sets.

The neoclassical growth theory distinguishes between *unconditional* (sometimes also called *absolute*) convergence and *conditional* convergence. The former assumes that all concerned economies have the same steady state, so that convergence will lead all the economies to a common (per capita) income level. On the other hand, the concept of conditional convergence assumes that the economies concerned have different steady states, and convergence will lead these economies to their respective steady state levels of per capita income. An important issue in studying convergence is therefore to ascertain whether convergence should be thought to be unconditional or conditional and, in the latter case, how to condition adequately for potential differences in steady state per capita income.

One appealing aspect of regional convergence study is that, while convergence across countries is generally thought to be only conditional, regional convergence can be assumed to be unconditional. This is because it is more plausible to argue that different regions of the same country have the same steady state per capita income level than it is to argue the same for different countries. In studying regional convergence, therefore one may be less concerned about controlling for differences in steady state. This is no mean advantage, because controlling for steady state differences adequately is not an easy task.

However, even in studying regional convergence, many researchers have thought it necessary to allow for some differences in the steady state. This amounts to treating convergence across regions as also a process of conditional convergence, at least to a certain extent. For example, in their study of convergence across the US states, Barro and Sala-i-Martin (1991, 1992a) allow for an industry composition index that varies across the states. The same has been the case with many other studies of regional convergence. Thus the study of regional convergence does not automatically imply study of unconditional convergence.

Of particular interest with regard to regional convergence is the unfolding experience of China. Whether or not convergence is taking place across Chinese provinces is of much significance because of several reasons. First is the sheer number of people whose welfare is involved with the issue. The population size of most of the Chinese provinces is larger than that of many countries of the world. Second, China's recent economic growth is taking place under a unique institutional setting in which a Communist Party-led government, that formerly practiced central planning, is now implementing market reforms. Reducing regional disparity was, and still is, a professed political goal of the Communist Party-led government. Under central planning, commands and administrative means were mainly used to achieve this political goal. A market economy, on the other hand, relies on spontaneous forces of the market. Technically, the convergence implication of the neoclassical growth theory hinges on diminishing returns to capital accumulation and diminishing opportunities for technological diffusion as an economy approaches the global technology frontier. It does not depend on institutional arrangements under which capital accumulation and technological diffusion take place. In principle, convergence can therefore occur under both central planning and market. The Chinese experience may therefore help to see whether the switch from central planning to market mechanism is making any difference with regard to convergence. It is from this viewpoint that the analysis of the Chinese convergence experience is of additional interest.

Several earlier papers have examined the issue of convergence across Chinese provinces. These include Lyons (1991), Chen and Fleisher (1996), Jian, Sachs, and Warner (1996), Gundlach (1997), Makino (1997), Raiser (1998), Kawabata and Meng (2000), Yao and Zhang (2001), Zhang (2001), Bao et al. (2002), Brun et al. (2002), and Cai et al. (2002). Of these, Lyons (1991) is a pioneering study that examines regional disparity using newly available provincial income accounts data. Focusing on coefficient of variation of per capita income across provinces,[5] Lyons finds that disparity across provinces increased during 1966-76, the period of 'Mao radicalism,' and decreased during the 'early years of liberalization' (1962-65) and during the first stage of economic reform (1978-83). Lyons however does not try to relate his study to any theory of convergence.

Jian, Sachs, and Warner (1996), on the other hand, proceed from the neoclassical convergence paradigm and use the concepts of both β- and σ-convergence to investigate the issue. (We discuss these concepts in more detail in the next section.) They analyze per capita GDP data of 15 provinces for 1952-78 and of 28 provinces for 1978-92. The authors find evidence of convergence during the immediate post reform period of 1978-85. However, they report that convergence gave way to divergence since 1985.

In studying regional convergence in China, researchers have often included additional right hand side variables in their regressions. Table 1 provides a list showing the variables that have been used by different authors. Inclusion of these variables would imply that the convergence investigated in these papers was deemed conditional.

---

[5] Lyons (1991) therefore examines what is now called σ-convergence.

Table 1: List of conditional variables in Chinese convergence studies

Bao et al. (2002)	Distance to coastline, Coastline length, Population distribution
Brun et al. (2002)	Population growth, Population, Inflation rate, FDI, SOE
Cai et al. (2002)	Human capital, Labor Participation, Labor productivity, Investment, Marketization
Chen and Fleisher (1996)	Coastal location
Jian, Sachs, and Warner (1996)	Agriculture share, Coastal location
Raiser (1998)	Light industry, Investment rates
Yao and Zhang (2001)	Distance to growth center, Export
Zhang (2001)	Non-collective sector, Export, Coast and west dummy

From the point of view of the neoclassical convergence paradigm, these additional variables should be treated as determinants of the steady state. In actuality, however, researchers have not always been that keen on accurate interpretation of the additional variables. For example, Jian, Sachs, and Warner (1996) in their study include coastal dummy as an additional conditional variable in the regression. However, they are not very clear about how to interpret this variable and the impact of its inclusion on the results. According to their analysis, the observed convergence during 1978-90 was a result of rural coastal provinces growing faster as a result of policy advantages.[6]

Raiser (1998) by contrast provides a more perceptive analysis and tries to stay within or close to the neoclassical convergence paradigm. He concludes that instead of divergence, what has happened since 1985 is better characterized as a weakening of convergence. Raiser notes that the weakening of convergence could be the result of either (a) 'shifts' in the steady state of some relatively richer (coastal) provinces or (b) reduction in capital mobility. His paper shows that capital mobility did not slacken in post-reform years. Instead, capital is directed to *richer* interior provinces (because of political considerations) rather than to poorer interior provinces. Thus, capital mobility is having a perverse effect so far as convergence is

[6] The confusing nature of Jian, Sachs, and Warner's (1996) analysis has been noted by other researchers too. Raiser (1998, p. 9), for example, notes that "…once the authors (Jian, Sachs, and Warner) control for the two factors (namely 'agriculture share' and 'coastal'), the evidence for convergence surprisingly disappears altogether. This makes an interpretation of their results difficult in the light of the foregoing theoretical analysis."

concerned. Instead of convergence among interior provinces, the misdirected capital mobility is causing divergence among them. Alongside, according to Raiser, the steady state income level of the coastal provinces underwent a positive shift due to the change in Chinese reform's focus from agriculture to industry beginning 1984. Raiser proves this claim by allowing a 'coastal' dummy both in intercept and interaction form in the growth-convergence regression and finding them to be significant. The dummy however does not reveal the actual structural characteristics that lie beneath it. Raiser therefore offers additional analysis in order to explain the dummy (or the shift). He thinks that the steady state income of coastal provinces could experience a positive shift, because they had two structural advantages allowing them to benefit more from the industrial reforms introduced in 1984. These two advantages were: (a) greater incidence of light industry and (b) greater prevalence of non-state enterprises (NSE). To test these assertions, Raiser runs regressions including ratio of light industry to total and the ratio of NSEs to total. Though the latter does not prove significant, the former does. Overall Raiser concludes that the post-1984 convergence slowdown was the combined effect of reversal of convergence among interior provinces, on one hand, and growing regional disparity between interior and coastal provinces, on the other.[7]

Raiser's analysis therefore shows that changes in income distribution across Chinese provinces are quite complex. Instead of a single overall pattern, there are different sub-patterns occurring in different parts. It is difficult to capture this complexity by simple summary indicators provided by values of either σ or β. Raiser tries to unearth this

---

[7] "Apparently, the convergence slow-down has been the result of a combination of growing income disparities between regions and a partial reversal of convergence within the interior region." (Raiser 1998, p. 13) In a sense, despite a different, more theoretically satisfying interpretation, Raiser's (1998) answer is similar to that of Jian, Sachs, and others. Raiser himself recognizes this: "My story is similar to the accounts in Yussuf (1994) and Jian, Sachs, and Warner (1996) in that it stresses rural industrialization and access to export markets." (p. 13) Raiser however adds that "in contrast to the latter source, I do find conditional convergence across all provinces once I control for a possible shift in the coastal steady state." (p. 13) Thus Jian, Sachs, and Warner (1996) and also to some extent Raiser (1998) seem to suggest that agricultural reforms initiated in 1978 benefited Chinese coastal provinces more, because these were agricultural. Since these were also relatively poor, the process helped convergence. However, by the time industrial reforms were introduced in 1984, these provinces have become relatively rich (because of earlier agricultural reforms), and yet they could benefit more from the new industrial reforms, causing divergence. This story however does not conform to the impression that one gets from other articles, which suggest that it were the interior provinces that benefited more from the agricultural reforms. For example, Cai, Wang, and Du (2002) seem to suggest that 1978 agricultural reforms actually benefited the interior provinces more. According to these authors, "The household responsibility system began in Anhui and Sichuan provinces, and experimental reform of SOEs, aimed at enlarging autonomy while sharing profits with the state, also began in the western regions (e.g. Sichuan province).…At the time, agricultural output counted for nearly one-third of total GDP in the country and even more in the central and western regions. Because the initial stages of reform particularly benefited agricultural development, which was relatively important in the central and western regions where productivity was very low, it reduced regional disparity." This is a very different story, and it is necessary to find out which of these two stories is more valid.

complexity as much as possible staying within the regression methodology and relying on the concept of β-convergence. However, his own analysis shows that the underlying changes in the shape of the distribution is far more complex than can be captured by estimated values of β. A proper understanding of regional convergence and evolution of regional disparity in China requires knowledge of the entire shape of the distribution and the changes that this shape is undergoing. It therefore requires a different methodological approach.

3. Methodological Issues

From the point of view of methodology, convergence research has proceeded along several approaches.[8] One of these focuses on what is called the β-convergence. This approach consists mainly of running regression (either cross-section or panel) with growth as the dependent variable and the initial income level as the right hand side variable, either in addition to other conditioning variables (in case of *conditional convergence*) or without any additional conditional variables (in case of *unconditional* or *absolute convergence*). In this regression framework, β, the coefficient of the initial income variable, is taken as the evidence for convergence, with a significant negative value denoting presence of convergence, an insignificant value denoting absence of convergence, and a significant positive value denoting divergence.

The concept of β-convergence and the associated regression methodology have been subjected to considerable criticism. A large part of this criticism focuses on estimation problems of β. Whether or not β can be correctly estimated is a thorny econometric issue, and many researchers continue to address this issue. However, there are two more fundamental criticisms. One of these relates to Galton Fallacy. Friedman (1994) and others[9] have pointed to the fact that a negative value of β does not necessarily imply a reduction in the dispersion (which is what convergence according to them should mean) of the distribution of per capita income across economies, and hence by attributing such an implication, researchers are falling into Galton Fallacy. The merit of this criticism has since been recognized by the literature, and it is now acknowledged that a negative β can only be a *necessary* pre-condition for convergence but not a *sufficient* condition. The Galton Fallacy criticism of β-convergence suggests that it would be better to compute directly some

---

8 See Islam (2003) for a discussion of different methodological approaches to convergence study.
9 See Quah (1993a)

measures of dispersion, such as the standard deviation (σ), etc. in order to examine whether the dispersion is decreasing over time. Convergence judged by changes in σ (hence known as the σ-convergence) would satisfy critiques worried about Galton Fallacy. However, a deeper analysis shows that the correspondence between decrease/increase in the value of σ and presence/absence of convergence is not that simple.[10]

The second line of criticism of β-convergence goes further, and it extends to σ-convergence too. This criticism, led by Danny Quah (1993b, 1996a), emphasizes the fact that both β and σ provide only summary statistics of the distribution of interest and hence are not adequate for understanding what is happening to the distribution. For example, the coefficient β pertains to the *mean* of the relevant conditional distribution. Similarly, σ provides a summary statistic of dispersion. Underneath the same changes in σ, there may be very different changes in the distribution. Hence by just looking at β or σ it is not possible to get an adequate understanding of the distribution dynamics, and in fact it is possible to arrive at erroneous conclusions. Quah therefore advocates convergence study using the distribution approach. This approach takes the entire shape of the distribution as the object of study and strives to estimate the distribution and reveal the changes in its shape.

Most of the research on convergence across both countries and regions has however used the concept of β-convergence and the associated regression methodology, despite the shortcomings mentioned above. This has been the case with studies of regional convergence in China too. One reason for the popularity of the β-convergence approach is the ease with which it can be implemented. In these days of ubiquitous software packages, regressions are easy to run, and the regression methodology also makes it easy to switch between investigations of unconditional and conditional convergence. By contrast, revealing the shape of the distribution poses a far more complex statistical task. That is why application of the distribution approach (beyond computation of the standard deviation, σ) has been limited. As far as our knowledge goes, none of the previous papers has used the distribution approach to study regional convergence in China. This paper therefore represents a maiden attempt in this regard.[11]

---

[10] See Barro and Sala-i-Martin (1995, 2004) and Islam (2003) for more details regarding the relationship between β and σ convergence.

[11] Sakamoto (2001) represents an earlier beginning of this research.

4. Markov Transition Matrix Methodology

In the growth and convergence literature the distribution approach has been pursued almost single-handedly by Danny Quah. In a series of papers, Quah (1993b, 1996a, 1996b) has developed the methodology for implementation of this approach. The methodology is based on the use of Markov (probability) transition matrix to model the change in distribution from one period to the next. In the following we present the essentials of this methodology before turning to its implementation for the Chinese case.[12]

Let $n \times 1$ vector F_t gives the distribution at time *t*, with *n* being the number of states distinguished to represent the distribution. In case of income distribution, as is in this paper, each state represents an income interval. Let *M* be the (*n* by *n*) Markov transition matrix governing the transformation of F_t into F_{t+1}, the distribution for *t+1*, so that we have

(1) $$F_{t+1} = M' \cdot F_t .$$

For example, if *n* equals 3, the Markov matrix assumes the following form.

(2) $$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

with each element of the matrix, a_{jk}, giving the probability of transition from state j during the initial period to state k during the next. These elements are therefore referred to as Markov transition probabilities.

Assuming that the Markov transition probabilities remain the same over time (in other words, assuming that the Markov transition matrix remains unchanged), the distribution after *s* periods can be obtained by repeating equation (1) *s* number of times, so that we have:

(3) $$F_{t+s} = (M^s)' \cdot F_t .$$

Under certain regularity conditions on the elements of the Markov transition matrix *M*, as *s* goes to infinity, the distribution converges to an ergodic distribution, sometimes also referred to as the steady state distribution, $\overline{F}$.[13] In other words,

[12] See Durlauf and Quah (1999) for more details.
[13] Not all transition matrices yield ergodic or steady state distributions. For a transition matrix to produce ergodic distribution, one of the eigenvalues has to equal 1, while the remaining eigenvalues have to be below 1

(4) $$F_{t+s} = (M^s)' \cdot F_t \xrightarrow{s \to \infty} \overline{F} \ .$$

The ergodic or steady state distribution does not change, so that

(5) $$\overline{F} = M' \cdot \overline{F} \ .$$

Equation (4) shows that for a particular transition matrix *M*, it is possible to obtain a corresponding steady state or ergodic distribution. The ergodic distribution shows what the long run distribution is going to be like if the observed dynamics continue to hold.

Thus the Markov transition matrix methodology can be helpful in answering many unanswered questions regarding regional convergence in China. It may be used to find out what the recent dynamics are in the form of computed transition probabilities and what kind of distribution will evolve (i.e., the ergodic distribution) if these dynamics continue to hold. It is then possible to compare and see how the ergodic distribution differs from the initial and final (sample) period distributions. Such a comparison can provide a fuller picture of what is happening with regard to the regional convergence process.

5. Approximation of distribution in selected years using the Normal kernel

The analysis of this paper focuses on the period from 1952 to 2003. The regional time series data on per capita GDP are compiled from NBSC (various years) and Kato and Chen (2002) and are brought to constant 1978 prices. Let y_i denote per capita GDP of province *i* in 1978 prices, and $\bar{y}$ be the cross-section average of y_i. We first want to abstract from the shift in the mean of the distribution as reflected in the secular movement in $\bar{y}$. We therefore normalize the data from different years by their respective cross-section means, and take the log of the ratio of y_i to $\bar{y}$ as the variable to analyze. We denote this variable by x_i, so that

(6) $$x_i = \ln\left(\frac{y_i}{\bar{y}}\right) = (\ln y_i - \ln \bar{y}).$$

We begin by approximating the actual distribution of x_i for selected years using the normal kernel. (See Silverman 1986 for details). The density function used for the approximation is as follows:

---

in absolute value. Under these conditions, the transition matrix can be said to be an irreducible and non-cycle matrix.

(7) $$\tilde{f}(X)=\frac{1}{nh}\sum_{i=1}^{n}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\cdot\left(\frac{X-x_i}{h}\right)^2\right),$$

where x_i is an observed value of the variable, n is the number of sample observations (population weighted), and h is the window width (assumed to be 0.2). The range of x is assumed to lie between -1.50 and 2.30.

Figure 1 presents the approximated distribution for 1952, 1978, and 2003. As we can see, the shape of the distribution has changed considerably over time. Comparing the distributions for 1952 and 1978, we find that relatively more of the provinces find their per capita income to be less than the average in 1978 than was the case in 1952. However, the distributions for both these years seem to be unimodal. That is however not the case with the 2003 distribution, which displays a bi-modal characteristic. This change of the shape of the distribution from unimodal to bi-modal is something that neither β- nor σ-convergence research would able to decipher.

Figure 2 shows distributions for a few more years of possible interest, such as 1965 (beginning of the Cultural Revolution), 1992 (beginning of the third stage of reforms), 2000 (accession to WTO). The basic story remains the same. The across province distribution of per capita income in China has changed from a unimodal one to a bi-modal one over the sample period of 1952 to 2003. Having thus obtained an overview of the distributions and the changes in them, we now turn to modeling the dynamics of distribution using Markov transition matrices and producing corresponding ergodic distributions.

6. Distribution Dynamics and Ergodic Distributions

An important issue in modeling distribution dynamics using Markov transition matrix concerns discretization. It involves determining the number of states and the grid values to demarcate these states. In this paper we provide two alternatives. One is based on five states, and the other is based on seven states. Table 2 presents the results from five-grid analysis, while Table 3 does the same for the seven-grid analysis.

Figure 1

Approximation of actual distribution of $\ln y_i$ in 1952, 1978, and 2003

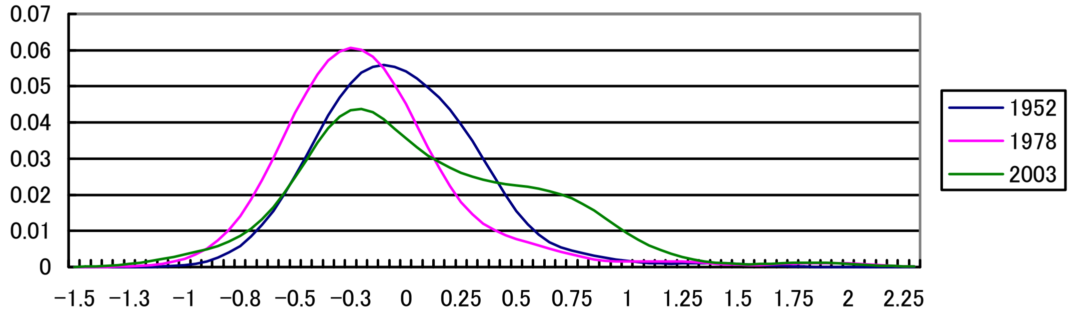
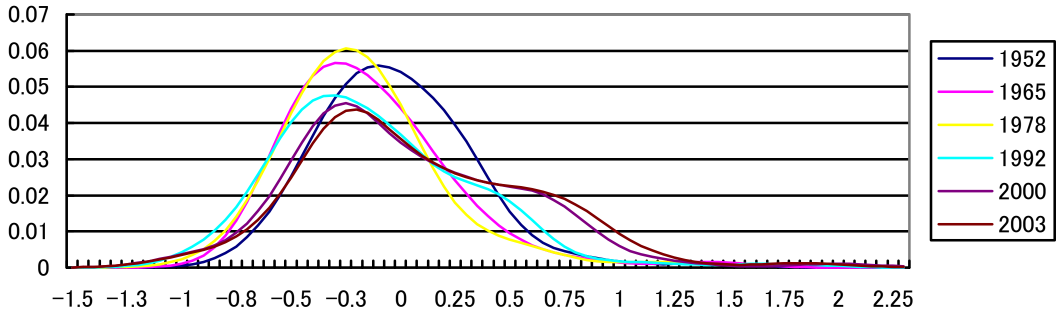


Figure 2

Approximation of actual distribution of $\ln y_i$ in particular years



With regard to the grid values to demarcate the states, there are several possibilities. This paper chooses grid values in a way so that overall (i.e., in the entire sample) the distribution of the actual values prove to be close to being uniform.[14] The number of provinces is 29 for the 1952-1978 period and 31 for the 1978-2003 period. All together therefore there are 1529

[14] Choosing gridlines to make the distribution uniform is a popular option. For example, Quah (1996a) too uses this option in his analysis of the U.S states. Another option is to choose the gridlines by fixed length, say by number of standard deviation or some other chosen interval (in order to produce equally spaced points on the range of the variable). In his analysis of 119 countries, Quah (1993, 1996a, and 1996b) opts for such arbitrarily chosen gridlines. One consideration guiding the choice of gridlines is the total number of transition episodes available in the data. If this number is large, the option of arbitrary gridlines becomes feasible, because all the states are likely to have enough number of transition episodes to make the analysis successful. In convergence studies, this is the case with a global sample of countries, which usually has a cross-section dimension of one hundred or more. The problem arises when the cross-section dimension is small, so that arbitrary gridlines may result in states with no or very few corresponding transition episodes, making the analysis infeasible or problematic. This is the case with the US states, where the cross-section dimension is less than fifty. In the case of China, the cross-section dimension is even smaller, only about thirty. That is why choosing gridlines so as to produce a uniform distribution proves to the better and feasible option in analyzing Chinese convergence.

observations on *annual* transition. A uniform distribution with five states would therefore have about 306 observations in each state. As can be seen from Table 2, the grid values (of x_i) prove to be -0.3486, -0.1970, -0.0309, and 0.2913, for the five-grid analysis.

The first panel of Table 2 shows the Markov transition matrix based on annual transition data from 1952 to 2003. The numbers in parentheses in the first column show the number of observations in the entire sample whose initial state belongs to the state represented by the respective row. Comparing these numbers across the rows (states) we see that the chosen grid points indeed resulted in a fairly uniform distribution in the entire sample. As already explained, the numbers in the cell represent Markov transition probabilities, so that a_{jk} gives the probability of transition from state j during the initial period to state k during the next period. The diagonal elements, a_{jj}, show the probability of a state remaining unchanged. Looking at the computed probabilities, we see that the diagonal values are higher at the two ends of the distribution than in the middle, indicating that the tendency toward persistence is high at the two ends of the distribution. By contrast, there is more flux in the middle. Provinces starting off from these middle states have higher probability to end up in a different state in the next period. The values of the matrix show that the probability of moving to a lower or higher state is almost the same for provinces starting off from the second lowest state. On the other hand, the probability of moving to a lower state is greater than that of moving to a higher state for provinces starting off from the middle and second highest state.

The numbers in the last row of the panel show the ergodic distribution that would result from the Markov transition matrix shown in the panel. We see that the probabilities at the two ends exceed the probabilities in the middle. There is indeed some thinning out of the middle, and as a result the distribution has attained a bi-modal characteristic. This agrees with the approximated distribution for 2003 presented in Figures 1 and 2. As we noticed, that distribution for 2003 displays bi-modality.

The second panel of Table 2 presents analogous results for the pre-reform period of 1952 to 1978. Looking at the transition probability matrix for this period, we see again that the diagonal values for the lowest and topmost states prove to be higher than those for the states in between, indicating again that persistence at the extremes is more pronounced. However, for this period the downward probabilities for the third and fourth state far outweigh the upward probabilities, so much so that the resulting ergodic distribution proves to be bottom heavy, as can be seen from the values in the last row of the second panel. This again agrees with the approximated distribution for 1978 presented in Figure 1. We saw that distribution to be more to the left of the distribution for 1952.

The third panel of Table 2 shows analogous results for the post reform period of 1979-2003. The dynamics of this period seem to be of opposite nature than those of the pre-reform period, and as a result, the ergodic distribution proves to be top-heavy, as can be seen from the values in the last row. The results of the third panel therefore provide ground for optimism regarding Chinese regional convergence. It shows that if the post-reform dynamics persist, more of the Chinese provinces will move to the higher income groups, with relatively few remaining in the low income ones.

The actual period of 1952-2003 embodies opposing influences. On the one hand, there is the downward dragging influence of the pre-reform period. On the other hand, it has the upward-lifting influence of the post reform period. The approximated actual distribution of 2003 and also the ergodic distribution based on the combined dynamics of the two sub-periods therefore produce a bi-modal distribution.

Table 3 shows analogous results with a finer grid division. We now have seven states, corresponding to the grid values -0.4090, -0.2873, -0.1711, -0.0589, 0.1110, and 0.5088. The number of observations in the first column of the Table shows that these grid values do produce a distribution that is close to being uniform. The first panel of the Table shows the Markov transition matrix based on annual transition data for the entire sample period of 1952-2003. The second and the third panel show that same for the pre-reform period 1952-1978 and post-reform period 1978-2003, respectively.

We see that the results prove qualitatively similar to those obtained from the analysis based on five states, presented in Table 2. The dynamics contained in the transition matrix based on the data for the entire sample produce a bi-modal ergodic distribution, as can be seen from the last row of the first panel. The dynamics contained in the transition matrix based on the pre-reform period produces a bottom-heavy ergodic distribution. Finally, the dynamics contained in the transition matrix based on the post-reform period yields a top-heavy ergodic distribution. This congruence of results based on two alternative discretizations (with regard to the number states) lends certain robustness to the results of the analysis.

The dynamics captured by the computed Markov transition matrix depends to a great extent on the period over which the transition is considered. The results presented in Tables 2 and 3 are based on annual transition, though averaged over fairly long periods of 51, 26, and 25 years, respectively, in the first, second, and third panel of these Tables. However, it may be argued that growth issues are essentially of long term nature, and hence annual transitions are not the right dynamics to depend on in examining growth regularities.

To address this concern, we next extend the analysis to five-year transition. Obviously, we now have fewer transition episodes. For the entire sample period of 1952-2003, we have 10 transitions.[15] With about thirty provinces, we have about 300 transitions, and with five grids, we expect to have about 60 transitions for each state if the distribution is uniform. We keep the gridlines the same as before, and interestingly they do produce strikingly uniform distributions of transitions in the sample data.

Table 4 presents the results from a five-grid analysis based on five-year transition. As before, the first panel contains the results for the entire sample period of 1952-2003, the second for the pre-reform period of 1952-1978, and the third for the reform period of 1978-2003. Clearly, when viewed over five years, more provinces undergo changes in their state than was the case when viewed over just one year. Thus the diagonal probability values now prove to be smaller than those in Table 2, showing more provinces experiencing changes in their state in either direction. Correspondingly, the off diagonal probability values are now larger.

However, in qualitative terms the results from five-year transition prove similar to those we obtained earlier on the basis of one-year transition. In each case, we find more persistence at the extremes and more movement in the middle. The ergodic distribution on the basis of the dynamics of the entire sample period displays a bi-modal characteristic. The ergodic distribution based on the dynamics of the pre reform period proves to be bottom heavy. Finally the ergodic distribution on the basis of the dynamics of the reform period proves to be top heavy. So even though the dynamics considered over the entire period produce bi-modal distribution, the dynamics contained in the reform period suggest that more Chinese provinces will catch up with richer provinces, and hence prove more encouraging from the point of view of convergence.

7. Concluding Remarks

This paper investigates the process of convergence across Chinese provinces using the distribution approach. Under this approach Markov transition matrix is used to capture the transition dynamics and produce corresponding ergodic distributions. This helps to see where

[15] The total time span is of 51 years. We take one transition to be of six years in order to cover the entire time span.

the distribution is headed if the current dynamics hold. The distribution approach therefore allows checking what is happening to the entire shape of the distribution, instead of knowing only what is happening to its mean (as in β-convergence studies) and dispersion (as in σ-convergence studies).

The analysis shows that the distribution of per capita income across Chinese provinces has attained a bi-modal characteristic over time. This conforms to the general impression of China observers that disparity has increased across Chinese provinces and that instead of convergence, a process of divergence is at work. However, examination of the dynamics of pre- and post-reform periods separately indicates that the observed bi-modality is a combined product of very different type of dynamics that prevailed during these two sub-periods. The ergodic distribution on the basis of the recent, post-reform period dynamics proves to be negatively skewed, with more provinces at the higher end of the distribution. This suggests that if the recent period dynamics hold, more Chinese provinces will catch up with the richer ones. Thus there are grounds to be more optimistic about regional convergence in China.

The distribution approach based on the Markov transition matrix, while having advantages in many respects, has its own limitations. The analysis involves discretization of the distribution and is therefore not entirely free from the influences of the options that exists in this regard. Also, the stochastic process may be better described by higher order Markov process instead of the first order process, as used in this paper (and in most other growth studies). Implementation of higher order processes is however not easy, given the limited number of transition episodes that are generally available in the data. Finally, the analysis is by and large a type of reduced form analysis, without having direct structural underpinnings. For example, most of the convergence analysis using distribution approach has been implicitly analysis of unconditional convergence. While it is in principle possible to use the framework to analyze conditional convergence by pre-filtering the data to take out the influence of the conditioning variables, this has generally not been done. Future research can be directed at overcoming these limitations.

Table 2: Transition matrix based on one-year transition

(5 grids)

1952-2003		Grid line (highest point)				
		-0.3486	-0.1970	-0.0309	0.2913	Inf
Samples	(309)	0.8738	0.1262	0	0	0
	(304)	0.1250	0.7533	0.1184	0.0033	0
	(307)	0.0033	0.1270	0.7687	0.1010	0
	(306)	0	0.0033	0.1046	0.8497	0.0425
	(303)	0	0	0	0.0363	0.9637
	Ergodic	0.2093	0.2064	0.1882	0.1825	0.2136
1952-1978		Grid line				
		-0.3486	-0.1970	-0.0309	0.2913	Inf
Samples	(145)	0.8345	0.1655	0	0	0
	(153)	0.1765	0.6536	0.1634	0.0065	0
	(142)	0.0070	0.1972	0.6831	0.1127	0
	(168)	0	0.0060	0.1310	0.8155	0.0476
	(146)	0	0	0	0.0685	0.9315
	Ergodic	0.2669	0.2424	0.1967	0.1734	0.1206
1978-2003		Grid line				
		-0.3486	-0.1970	-0.0309	0.2913	Inf
Samples	(164)	0.9085	0.0915	0	0	0
	(151)	0.0728	0.8543	0.0728	0	0
	(165)	0	0.0667	0.8424	0.0909	0
	(138)	0	0	0.0725	0.8913	0.0362
	(157)	0	0	0	0.0064	0.9936
	Ergodic	0.0661	0.0829	0.0906	0.1137	0.6467

Table 3: Transition matrix based on one-year transition

(7 grids)

1952-2003		Grid line (highest point)						
		-0.4090	-0.2873	-0.1711	-0.0589	0.1110	0.5088	Inf
Samples	(220)	0.8818	0.1045	0.0136	0	0	0	0
	(219)	0.1142	0.7352	0.1507	0	0	0	0
	(218)	0.0046	0.1468	0.6789	0.1560	0.0138	0	0
	(220)	0	0.0182	0.1500	0.7000	0.1273	0.0045	0
	(221)	0	0	0.0136	0.1357	0.7376	0.1131	0
	(216)	0	0	0	0.0046	0.1111	0.8380	0.0463
	(215)	0	0	0	0	0	0.0326	0.9674
	Ergodic	0.1453	0.1447	0.1412	0.1324	0.1259	0.1282	0.1823
1952-1978		Grid line						
		-0.4090	-0.2873	-0.1711	-0.0589	0.1110	0.5088	Inf
Samples	(109)	0.8532	0.1193	0.0275	0	0	0	0
	(94)	0.1809	0.6170	0.2021	0	0	0	0
	(116)	0.0086	0.1810	0.6034	0.1810	0.0259	0	0
	(99)	0	0.0404	0.2323	0.5758	0.1515	0	0
	(115)	0	0	0.0261	0.1565	0.6783	0.1391	0
	(120)	0	0	0	0.0083	0.1667	0.7833	0.0417
	(101)	0	0	0	0	0	0.0495	0.9505
	Ergodic	0.2158	0.1665	0.1819	0.1260	0.1257	0.1000	0.0841
1978-2003		Grid line						
		-0.4090	-0.2873	-0.1711	-0.0589	0.1110	0.5088	Inf
Samples	(111)	0.9099	0.0901	0	0	0	0	0
	(125)	0.0640	0.8240	0.1120	0	0	0	0
	(102)	0	0.1078	0.7647	0.1275	0	0	0
	(121)	0	0	0.0826	0.8017	0.1074	0.0083	0
	(106)	0	0	0	0.1132	0.8019	0.0849	0
	(96)	0	0	0	0	0.0417	0.9063	0.0521
	(114)	0	0	0	0	0	0.0175	0.9825
	Ergodic	0.0347	0.0488	0.0507	0.0782	0.0799	0.1783	0.5294

Table 4: Transition matrix based on five-year transition
(5 grids)

1952-2003		Grid line (highest point)				
		-0.3486	-0.1970	-0.0309	0.2913	Inf
Samples	(60)	0.7000	0.3000	0	0	0
	(62)	0.2931	0.5172	0.1724	0.0172	0
	(60)	0.0161	0.1774	0.5968	0.2097	0
	(61)	0	0.0476	0.1905	0.6349	0.1270
	(59)	0	0	0	0.1053	0.8947
	Ergodic	0.2262	0.2219	0.1755	0.1706	0.2058
1952-1978		Grid line				
		-0.3486	-0.1970	-0.0309	0.2913	Inf
Samples	(30)	0.7692	0.2308	0	0	0
	(29)	0.3103	0.4483	0.2069	0.0345	0
	(30)	0.0370	0.2963	0.5185	0.1481	0
	(33)	0	0.0556	0.2500	0.6111	0.0833
	(25)	0	0	0	0.1852	0.8148
	Ergodic	0.3892	0.2685	0.1752	0.1153	0.0519
1978-2003		Grid line				
		-0.3486	-0.1970	-0.0309	0.2913	Inf
Samples	(30)	0.6471	0.3529	0	0	0
	(33)	0.2759	0.5862	0.1379	0	0
	(30)	0	0.0857	0.6571	0.2571	0
	(28)	0	0.0370	0.1111	0.6667	0.1852
	(34)	0	0	0	0.0333	0.9667
	Ergodic	0.0594	0.0760	0.0699	0.1212	0.6736

References

Bao, S. M., G. H. Chang., J. D. Sachs., and W.T. Woo. (2002) "Geographic factors and China's regional development under market reforms 1978-1998," *China Economic Review,* 13, pp.89-111.

Barro, R. J. and X. Sala-i-Martin. (1991) "Convergence across States and Regions," *Brookings Papers on Economic Activity*, No. 1, pp. 107-182

Barro, R. J. and X. Sala-i-Martin. (1992a) "Convergence," *Journal of Political Economy* 100(2) 233-251

Barro, R. J. and X. Sala-i-Martin. (1992b) "Regional Growth and Migration: A Japan-United States Comparison, *Journal of the Japanese and International Economies*, 6 (December), 312-346

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

Barro, R. J. and X. Sala-i-Martin (2004) *Economic growth* (Second edition), Cambridge: MIT Press

Brun, J. F., J. L. Combes, and M. F. Renard. (2002) "Are there spillover effects between coastal and non coastal regions in China?" *China Economic Review,* 13, pp.161-169

Cai, F., D. Wang., and Y. Du. (2002) "Regional disparity and economic growth in China: The impact of labor market distortions," *China Economic Review,* 13, pp.197-212.

Cashin, Paul and Ratna Sahay (1995), "Internal Migration, Center-State Grants, and Economic Growth in the States of India," IMF Working Paper, Washington DC.

Chen, J. and B.M. Fleisher. (1996) "Regional income inequality and economic growth in China," *Journal of Comparative Economics,* 22(2), pp. 141-164.

Coulombe, Serge and Frank C. Lee (1993), *Regional Income Disparities in Canada*, Unpublished Ph.D. dissertation, University of Toronto

Durlauf, Steven N. and Danny Quah (1999), "The New Empirics of Economic Growth," in John B. Taylor and Michael Woodford (editors), *Handbook of Macroeconomics*, Vol. 1, Amsterdam: North Holland

Friedman Milton (1994), "Do Old Fallacies Ever Die?" *Journal of Economic Literature,* 30 (December, No. 4), pp. 2129-2132.

Gundlach, Eric (1997), "Regional Convergence of Output per Worker in China: A Neoclassical Interpretation," *Asian Economic Journal*, Vol. 11 (No. 4), pp. 423-442

Holz-Eakin, Douglas (1993), "Solow and States: Capital Accumulation, Productivity, and Economic Growth," *National Tax Journal*, Vol. 46, pp. 425-439

Hossain, Akhtar (2000), "Convergence of per capita output levels across regions of Bangladesh: 1982-97," IMF Working Paper, Washington DC.

Islam, Nazrul. (2003) “What Have We Learnt from the Convergence Debate? A Review of the Convergence Literature,” *Journal of Economic Surveys*, 17 (No. 3, June), pp. 309-62

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

Kato, H. and G. H.Chen. (2002), *China: Vol. 12 of Long term economic statistics of East Asia*, edited by Asia information center, Takushoku University (in Japanese), *Keiso shobou*

Kawabata, K. and J. J. Meng (2000), “Convergence of regional economy in China: Statistical verification by cross section and time series analysis (in Japanese),” *Asia Keizai,* 41(6), pp. 20-33

Lyons, T. P. (1991) “Inter-provincial disparities in China: Output and consumption 1952-1987,” *Economic Development and Cultural Change,* 39(3), pp. 471-506

Makino, M. (1997) “Inter-regional disparities in China: Welfare vs. productivity,” *Osaka City University Economic Review,* 32(1and 2), pp. 1-36

National Bureau of Statistics of China (NBSC) (1997), “The gross domestic product of China (1952-1995),” *Dongbei University of Finance and Economics press*

National Bureau of Statistics of China (NBSC) (1998), *China regional economy: A profile of 17 years of reform and opening-up*, Beijing: China statistics press

National Bureau of Statistics of China (NBSC) (various years), *China Statistical Yearbook*,” Beijing: *China statistics press*

Persson, Joakim (1997), “Convergence across the Swedish counties, 1911-1993,” *European Economic Review*, Vol. 41 (December), pp. 1835-1852

Quah, Danny (1993a), “Galton’s Fallacy and Tests of the Convergence Hypothesis,” *Scandinavian Journal of Economics,* 95 (4), pp. 427-443

Quah, Danny (1993b), “Empirical Cross-Section Dynamics in Economic Growth,” *European Economic Review,* 37, pp. 426-434

Quah, Danny (1996a), “Empirics for Economic Growth and Convergence,” *European Economic Review,* 40 (No. 6, June), pp. 1353-75

Quah, Danny (1996b), “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics,” *Economic Journal*, 106 (July), pp. 1045-1055

Raiser, M. (1998) “Subsidizing inequality: Economic reforms, fiscal transfers and convergence across Chinese provinces,” *Journal of Development Studies,* 34(3), pp. 1-26

Sakamoto, H. (2001) “Regional economic inequality and its convergence in China,” (in Japanese),” *Osaka Economic Papers,* 51(1), pp. 89-110

Sala-i-Martin, Xavier (1990), On Growth and States, Unpublished Ph.D. dissertation, Harvard University, Cambridge, MA

Sanchez-Robles, Blanca and Jose Villaverde (2001), “Polarizacion, Convergencia, y

Movilidad entre las provincias espanolas, 1955-1997," *Revista Asturiana de Economica*," May, 259-270

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, London, Chapman and Hall.

Temple, Jonathan (1999), "New Growth Evidence," *Journal of Economic Literature*, Vol. 37 (March), pp. 112-156

Yao, S. J. and Z. Y. Zhang (2001), "On regional inequality and diverging clubs: A case study of contemporary China," *Journal of Comparative Economics,* 29, pp. 466-484

Yusuf, S. (1994), "China's Collective and Private Enterprises: Growth and Its Financing," in G. Caprio (ed.), *Building Sound Finance in Transition Economies*, IMF, Washington DC.

Zhang, W. (2001) "Rethinking regional disparity in China," *Economics of Planning,* 34, pp. 113-138