Distribution Dynamics and Convergence among 75 Cities and Counties in Yangtze River Delta in China: 1990-2005

Hiroshi Sakamoto, ICSEAD
and
Jin Fan, Research Centre for Jiangsu Applied Economics, Jiangsu Administration Institute

Working Paper Series Vol. 2009-25
November 2009

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

Distribution Dynamics and Convergence among 75 Cities and Counties in Yangtze River Delta in China: 1990-2005

Hiroshi Sakamoto[♦]
Jin Fan[*]

Abstract

This paper applies the distribution dynamics method to study the per capita income disparity from 1990 to 2005 among the 75 cities and counties in the Yangtze River Delta (YRD). The main conclusions are as follows: First, the distribution of per capita income across YRD has changed from bi-modal to being positively skewed over the period 1990–2005; the income disparity has been reduced in the 8th Five-Year Plan, enlarged in the 9th Five-Year Plan and then reduced again somewhat in the 10th Five-Year Plan. Second, the main contribution to disparity comes from the intra disparity of the Jiangsu region; especially, the distribution density of Jiangsu is bi-modal over the period. Third, the rich cities and the poor cities developed independently and steadily at different speeds. Fourth, the cause of the convergence in the YRD is the decline of the rich cities and not the poor cities' progress. In the end, there is relative independence in the region's development, but long-term the YRD shows the weaker trend of convergence.

Key words: Yangtze River Delta; Income Distribution; Kernel Density; Markov Transition Matrix
JEL: C16, O15, R11

[♦] Research Associate Professor, The International Centre for the Study of East Asian Development (ICSEAD), Kitakyushu 11-4 Otemachi, Kokurakita, Kitakyushu, 803-0814 Japan
Tel: +81 93 583 6202; Fax: +81 93 583 4602; E-mail:sakamoto@icsead.or.jp
The authors would like to thank for their useful comments the participants of the second China-Japan joint seminar of applied regional science where an earlier version of this paper was presented. Sakamoto's research was supported by grant number 21530247 financed by the Japanese Ministry of Education, Culture, Sports, Science and Technology. Fan's research was supported by grant number 70873052 financed by the Natural Science Foundation of China. Of course, the authors are entirely responsible for all the remaining errors.
[*] Professor, Research Centre for Jiangsu Applied Economics, Jiangsu Administration Institute
168 Jianye Road, Nanjing, 210004, China
Tel: +86 25 85013353; Fax: +86 25 84201784; E-mail: jfan@mx.cei.gov.cn

1. Introduction

Yangtze River Delta metropolitan area (YRD) is one of the world's six major metropolitan areas and in the "People's Republic of China national economic and social development in the 11th Five Year Plan," this metropolitan area was identified as a state-level regional economic circle. This not only reflects the region's economic strength and its tremendous achievements in economic development in the past, but also highlights its prominent position cannot be replaced in China's future economic development. One of the most prominent features in YRD is the development of the belonging cities. Shanghai, Jiangsu, and Zhejiang, these adjacent areas comprised 16 city groups in 2005. In comparing the cities' competitiveness, 15 of these are included in the top 50 most competitive cities. Also it has a large number of economic level, comprehensive strength and growth potential are very prominent county-level city in 2004, China's top 100 counties in the comparative rankings, the former top 10, in addition to the first, sixth, the remaining eight all belong to the counties in the YRD region; followed by Jiangsu's Kunshan City, Jiangyin City, Zhangjiagang City, Changshu City, Zhejiang's Xiaoshan District, the Jiangsu's Wujin District, Wujiang City, and Zhejiang's Shaoxing County. Although the metropolitan area, the YRD cities and city-county group, together contribute to the overall economic prosperity, there is some disparity individually. This paper will use the income distribution dynamic methods, namely the Gaussian kernel density estimates and Markov transition matrix, to illustrate the per capita GDP income shape and income distribution over time, of the 75 counties and cities of the YRD, and reveal their internal economic structure.

In the study of convergence, β convergence and σ convergence is discussed mainly from the point of view of methodology. However, β convergence and σ convergence provide only general information on convergence, and the information is often unable to explain the internal structure of economic development. The method of distribution as a test of the methods of economic convergence, including that between regions, can reveal the internal structure of the economy, as well as determine the dynamic development trends. Since Quah (1993, 1996a, b) proposed distribution methods, the technique has received more and more attention with increasing applications. In order to better describe the YRD economic structure and the internal dynamic movements, this paper will use the income distribution method of dynamic empirical study.

This paper is structured as follows: Section 2 presents the literature review. Section 3 introduces the data sources and related data processing. Section 4 estimates the income disparity using a traditional method. For understanding contribution of disparity at the regional level, it adapts decomposition analysis using Mean Log Deviation. Section 5 uses the Gaussian kernel density methods and estimates income distribution for 1990, 1995, 2000 and 2005 for the YRD metropolitan area using per capita GDP. Section 6 introduces the Markov transition matrix methods and transition matrix estimates as well as the convergence (ergodic) distribution. It also investigates the dynamic evolution of per capita GDP in the YRD metropolitan area. The last section offers some concluding remarks.

2. Literature Review

The convergence of economic growth refers to the negative correlation between the initial static indicators (per capita GDP or outputs) and the rate of economic growth. The three major tests of convergence hold common assumptions: (1) that the cross-section of time series data is compared assuming the homogeneity of static analysis (Abramovitz, 1956; Baumol, 1986; Barro and Sala-i-Martin, 1992; Wolff, 1991; Dowick and Nguyen, 1989); (2) that the cross-section of time series data can be compared when assuming the heterogeneity of static analysis (Barro and Sala-i-Martin, 1992; Mankiw, Romer and Weil, 1992); (3) that there is a dynamic panel data model (Islam, 1995; Lee, Pesaran and Smith, 1997).

With the above theoretical models, when using cross-section analysis of time series data, interpretation of the economic structure is based on fundamental analysis of the β coefficient and σ coefficient. Also, in academic research on China's regional convergence, the methods commonly used involve the β convergence and σ convergence. However, because income disparity across China's provinces is a very complex issue and not a single overall pattern, with different distribution patterns in different places, it is difficult to use β convergence and provide a simple σ convergence of information to capture this complexity (Raiser, 1998). There is, therefore, a need to find new ways to study various areas of the convergence issues in China. More and more researchers have begun applying the method of distribution dynamics in studying the issue of regional convergence in China.

Aziz et al. (2001) estimate the probability density of China's provincial-level income distribution for 1978–1997 based on the per capita GDP and they suggest that in the initial stage of reform, the income gap had declined, but overall China's inter-provincial income distribution is a bi-modal distribution of development. Xu and Li (2004) use the kernel density function method for 1978-1998 and estimate China's provincial-level income distribution based on the GDP per labor. Their findings agree with Aziz et al.'s (2001) that the formation of a single peak changes to twin peaks in the evolution of the trend, but their density map shows the trend more obviously than Aziz et al.'s (2001). Sakamoto and Islam (2008) study the convergence issues of China's provinces in the 1952–2003 period using the Gaussian kernel density distribution for selected years to describe the shape in the provinces' per capita GDP; they also use the Markov transition matrix to reveal the dynamic changes in distribution of shape, offering the most reasonable interpretation of the convergence issues of China's provinces thus far.

Many of the current studies of convergence concentrated mainly on the provincial level or between one province and its city, but not on an in-depth study of all counties within the urban area of convergence. In fact, the analysis of regional economic convergence, using a provincial or county spatial scale of the statistical data will reveal the economic disparities in different scales of space. County as a relatively complete basic space units and infrastructure level, to study can be more detailed information and in-depth reliable conclusion this is the analysis of the YRD evolution of the key. Therefore, this study has important practical significance and meaning for the model.

3. Data Sources and Compiling

This paper analyses the YRD region for the target years 1990–2005 evaluating the per capita GDP/GRP (Gross Regional Product) in 75 districts and counties. Sources of data are the "Statistical Yearbook of Shanghai" (1990–2005), "Jiangsu Provincial Statistical Yearbook" (1990–2005) and "Zhejiang Provincial Statistical Yearbook" (1990–2005).

In this paper, based on the 2006 Statistical Yearbook of regional standards, a study of 75 samples has been selected, including 8 cities and 36 counties (samples) in Jiangsu province,[1] 1 city (samples) in Shanghai[2], and 7 cities and 38 counties (samples) in Zhejiang province.[3] [4] For the 1990–2005 period, because of some administrative changes in the region, we provide the data's corresponding address (see Figure 1).[5]

The data of GRP is deflated by each provincial price index and based on the 2005 price. In addition, Shanghai's mega-cities have populations of over 10 million and their level of development was higher than that of the other samples, which could create a distortion in the kernel density distribution. Therefore, we consider population weighting for the following analysis.

4. Regional Income Disparity in YRD: Decomposition Analysis

---

[1] Detailed samples for Jinagsu are Nanjing (Urban, Lishui County, Gaochun County), Wuxi City (Urban, Jiangyin City, Yixing City), Changzhou City (Urban, Liyang City, Jintan City), Suzhou City (Urban, Changshu City, Zhangjiagang City, Kunshan City, Wujiang City, Taicang City), Nantong City (Urban, Hai'an County, Rudong, Qidong City, Rugao, Tong Zhou, Haimen), Yangzhou City (Urban, Baoying County, Yizheng City, Gaoyou City, Jiangdu), Zhenjiang City (Urban, Danyang City, Yangzhong City, Jurong City), Taizhou City (Urban, Xinghua City, Jingjiang City, Taixing City, Jiangyan city).

[2] There are 19 districts and counties in Shanghai city (Pudong New Area, Huangpu, Luwan District, Xuhui, Changning District, the Jingan District, Putuo, Zhabei District, Hongkou District, Yangpu District, Baoshan, Minhang and Jiading District, Jinshan, Songjiang District, Qingpu District, Nanhui District, Fengxian District and Chongming County in detail). However, because there are no income data for each, we assume one sample in this study.

[3] Detailed samples for Zhejiang are of Hangzhou (Urban, Jiande City, Tonglu County, Fuyang County, Linan City, Chun'an County), Ningbo (Urban, Yuyao City, Cixi City, Fenghua City, Xiangshan County, Ninghai County), Jiaxing City (Urban, Haining City, Pinghu City, Tongxiang City, Jiashan County, Haiyan), Huzhou City (Urban, Deqing County, Changxing County, Anji County), Shaoxing City (Urban, Zhuji City, Shangyu City, Shengzhou, Xinchang County), Zhoushan City (Urban, Daishan County, Shengsi Xian), Taizhou (Urban, waterfront City, Xianju County, the roof counties, three counties, Yuhuan County).

[4] According to the definition of YRD and its area, the north area of Jiangsu and the Southwest area of Zhejiang are not included in our dataset.

[5] Xiaoshan City is included in parts of Hangzhou. Nanjing's four northern districts and counties will be consolidated into two new districts. Wuxi revoked Xishan's city status making it two areas—Hui Shanqu and Xishan District—with Suzhou City incorporated into Wuxian.

First of all, the income disparity across regions in YRD is measured by using a traditional statistical technique. This is a necessary supplemental procedure, and its results are mutually related to the distribution dynamics methods that will be introduced later. In this paper, we introduce the Mean Log Deviation (MLD: the second Theil entropy) to show regional income disparity. This is a measurement method for considering differences in the population of each sample.

MLD L is the following specification for province k, sample i, population of each sample $N_{k,j}$ GRP of each sample $Y_{k,j}$.

$$L = \sum_k \sum_i \left(\frac{N_{k,i}}{N} \right) \cdot \log \left(\frac{N_{k,i}/N}{Y_{k,i}/Y} \right), \qquad (1)$$

where $N = \sum_k \sum_i N_{k,i}$ and $Y = \sum_k \sum_i Y_{k,i}$.

Next, we consider intra-MLD for province k (L_k)

$$L_k = \sum_i \left(\frac{N_{k.i}}{N_k} \right) \cdot \log \left(\frac{N_{k,i}/N_k}{Y_{k,i}/Y_k} \right). \qquad (2)$$

Then, L is decomposed by

$$\begin{aligned} L &= \sum_k \left(\frac{N_k}{N} \right) \cdot L_k + \sum_k \left(\frac{N_k}{N} \right) \cdot \log \left(\frac{N_k/N}{Y_k/Y} \right) \\ &= \sum_k \left(\frac{N_k}{N} \right) \cdot L_k + L_{BR} \end{aligned}, \qquad (3)$$

where L_{BR} is inter-MLD across province.

Figure 2 shows the trend of the income disparity in the YRD. It is found that the disparity was reduced once during the 8th Five Year Plan then rose again after that. However, it shrinks again after 2003 somewhat. The disparity has not necessarily changed greatly overall. The Figure also shows Jiangsu's contribution to disparity is the most and increasing when the decomposition of disparity is made among regions including Shanghai (L_{BR}), intra-region in Jiangsu ($L_{Jiangsu}$) and intra-region in Zhejiang ($L_{Zhejiang}$).[6] On the other hand, the contribution to disparity among regions (L_{BR}) has become small. It is found that the disparity of each sample in Jiangsu province in the YRD will have expanded though the YRD can be expected to register a reduction in the disparity at provincial level. Based on these results, the income distribution density in the YRD region will be measured.

5. Gaussian Kernel Density Distribution

Gaussian kernel density distribution is commonly used for estimating distribution. It is a non-parametrical estimation method. In this section, the change in the disparity of YRD is investigated by approximating the density function. The decomposition of the disparity using the MLD provides only summary information about the distribution. Therefore, for the detailed distribution change situation, it is more effective to examine the density function through estimation (Quah, 1996c, 1997; Sakamoto and Islam, 2008).

The approximation of distribution density is carried out as follows. Let X_i denote per capita GRP of sample i in 2005 prices, and $\bar{X}$ be the cross-section population weighted average of X_i. We first want to abstract from the shift in the mean of the distribution as reflected in the secular movement in $\bar{X}$. We therefore normalize the data from different years by their respective cross-section means, and take the log of the ratio of X_i to $\bar{X}$ as the variable for analysis. We denote this variable by Z_i, so that

$$Z_i = \ln\left(\frac{X_i}{\bar{X}}\right) = \ln X_i - \ln \bar{X}. \qquad (4)$$

[6] Because Shanghai has only one sample, we cannot calculate intra-regional disparity in Shanghai. Therefore, contribution to disparity at intra-regional level in Shanghai is zero. Contribution to disparity in Shanghai includes inter-regional disparity (L_{BR}).

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

$$\tilde{f}(Z) = \frac{1}{h}\sum_{i=1} \frac{w_i}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{Z - Z_i}{h} \right)^2 \right), \qquad (5)$$

where Z_i is an observed value of the variable, w_i is the population weight of sample i, and h is the window width. In this paper, it is assumed to be 0.25.[7] The range of Z is assumed to lie between -2.30 and 1.80.[8]

This paper chose the years 1990, 1995, 2000 and 2005 for evaluating per capita GRP level of 75 counties and cities in the YRD metropolitan area in the comparative study.[9] Moreover, we decomposed density for the three regions of Jiangsu, Shanghai, and Zhejiang.

The income distribution structure had changed greatly between 1990 and 1995 (see Figure 2). It changed from the bi-modal situation whose peaks were about -0.8 and 0.1 in relative values to an average mono-modal pattern. As a result, it seems that the income disparity experienced a reduction once. However, while the peak for 0.4 in relative value remains, distribution below 0.4 deteriorated gradually in 2000 and 2005. This is because the density at around -1.3 to -1.5 is quite poor as income level is increasing.

Comparing with Sakamoto and Islam (2008), income distribution density of the YRD is not bi-modal, but distribution keeps gradually expanding. The movement of income distribution density is somewhat similar to the statistical index in Figure 2.

Next, this income distribution structure is decomposed into three regions. Figure 4

---

7 According to Silverman's rule (Silverman, 1986), Gaussian kernel density is of the window width $h_n = 0.9AN^{-1/5}$. A is the standard deviation, N is sample number.

8 In actual measurement, measurement point Z is decided beforehand, and then value of the density function in relation to that is calculated. The measurement point was set at 0.05 intervals. Of course, the smaller this interval, the more accurate is the measurement.

9 The years 1990, 1995 and 2000 are the start years of "8th Five Year Plan," "9th Five Year Plan," and "10th Five Year Plan," respectively. Because of strong representation, this paper selected four years for comparative analysis.

shows the decomposition of three regions in 1990. The distribution structure of Jiangsu is clearly bi-modal which is one of the reasons the distribution in the YRD is bi-modal. On the other hand, Zhejiang is mono-modal but its peak exists between the two peaks of Jiangsu. Shanghai's higher income is reflected in a higher distribution than these two provinces.

Figure 5 shows the peak of income level in Shanghai is falling and that in Zhejiang is going up, while Jiangsu approaches its first peak (higher income peak) in 1995. Then, the density of the next peak (lower income peak) of Jiangsu is decreasing and the density below the lower peak is increasing inversely.

The result in 2000 is almost similar to that of 1995. The density below -1.0 rises gradually though the position of the peak income level has not changed that much. As a result, the entire distribution will have extended (see Figure 6). The position of the peak income level has still not changed in 2005, but density is decreasing in both Jiangsu and Zhejiang (below 0.08). This falling density is the result of poor distribution (see Figure 7).

The internal structure is complex though there seems to be a mono-modal pattern of distribution in the entire region. Especially, Jiangsu shows bi-modal distribution and the disparity is large as in Figure 2.[10] Therefore, it can be said that there is a difference in economic development at the prefecture level in the YRD. Further, the influence that this difference has on the future distribution will be estimated by using the Markov chain.

6. Approximation of Markov Transition Matrix and its Convergence Distribution

Next, the ergodic (future) characteristics of the income distribution structure will be examined by using the Markov chain. Quah (1993, 1996a, b) has developed the methodology for implementation of this approach. The methodology is based on the use of the Markov transition matrix to model the change in distribution from one period to the next. In the following we briefly present the essentials of this methodology.

[10] Of course, it is possible to be more specific by further decomposing this structure.

Let $n \times 1$ vector F_t give the distribution at time t, with n being the number of states distinguished to represent the distribution. In regard to income distribution, 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

$$F_{t+1} = M^t \cdot F_t . \qquad (6)$$

The Markov matrix assumes the following form.

$$M = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{j1} & \cdots & a_{jk} \end{pmatrix}, \qquad (7)$$

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 matrix remains unchanged, the distribution after several periods can be obtained by repeating equation (6) a number of times. And if repetitions go to infinity, the distribution converges to an ergodic distribution, sometimes also referred to as the steady state distribution, $\bar{F}$ .The ergodic or steady state distribution does not change, so that

$$\overline{F} = M^t \cdot \overline{F} . \qquad (8)$$

Equation (8) shows that for a particular transition matrix M, it is possible to obtain a corresponding steady state or ergodic distribution. Technically, the ergodic distribution is computed as the left eigenvector corresponding to the unit eigenvalue. The ergodic distribution shows what the long run distribution is going to be like if the observed dynamics continue to hold.

The transition probability matrix is estimated according to the following procedures.

First, after setting an appropriate grid line to all samples, the income state of each sample is defined. Second, the elements of the matrix are estimated by counting the change of the income state for each case, and dividing the total transition number by the total number of the income state before it changes.[11] In this procedure, the number of state changes is counted as one. However, a bias is expected in our dataset because there is great difference among the population share of our dataset.[12] Therefore, before calculating the total transition number, we apply the population weight to the number of state changes.[13] And, the element of the matrix is similarly calculated by the method of maximum likelihood.

An important issue in modeling distribution dynamics using the Markov transition matrix concerns discretization. It involves determining the number of states and the grid values to demarcate these states. With regard to the grid values to demarcate the states, there are several possibilities.[14] In this paper we report a four-grid (four states) analysis to make the sample in which the state change takes place for the period.[15] And, the grid line of income level is set by distributing all samples equally. In this case, income level is derived from equation (4) and is ordered from lowest income up. However, because a bias to the population size also remains in this case, the grid line of income level is set from the points of 25%, 50%, and 75% of the accumulation density (in total of all the densities for the period), using information on the density function estimated in the previous section. Therefore, the population weighted transition matrix is estimated using the above-mentioned two grid line methods.

Table 1 and Table 2 are results of the transition probability matrix in which the population weight was applied and the ergodic (convergence) distributions were indicated. The panel delimits for various periods at intervals of 5 years, 10 years, and 15 years. Table 1 set the grid line in the place where the sample had been equally distributed among four. As for the grid line, it is set at -0.710, -0.337 and -0.033 (using

[11] This is the same as the estimate by the method of maximum likelihood.
[12] Shanghai which has a population of more than 10 million is made one sample in our dataset.
[13] The population weight uses the simple arithmetic average of the population weight before the state changes and the population weight after the state changes.
[14] Therefore, the belief that this technique becomes arbitrary because of the researcher's subjectivity is incontrovertible.
[15] An appropriate ergodic distribution is not achieved when the state transition doesn't take place.

low income as the starting point), and includes 300 samples (75 samples × 16 years = 1,200 which is then divided by four) which have the income for each state distributed throughout the period. Mean value 0 was included for highest income state, and became an estimate after the bias had been somewhat corrected for.

The estimated value of the ergodic distribution for each period was described under each panel. In the estimate of 1990–2005 for all periods, the ergodic distribution became 0.1763, 0.1553, 0.1808, and 0.4876 measuring from the lowest income state. One finds that distribution was concentrated in the highest income state which included the mean value. Moreover, too much significance should not be attached to the slight possibility of bi-modality since the lowest income state has the third greatest density. It is noted that the distribution of the mono-modal pattern is shown in the ergodic distribution as well as in the estimate of the density function.

On the other hand, the concentration in the highest income state during 1990–1995 is shown (0.7046); however, it shows a tendency to being bi-modal in the next five years (0.2732, 0.2178, 0.1683, and 0.3407 measured from the lowest income state). Oppositely, the concentration in the lowest income state tends to be during 2000–2005 (0.7814). One finds that the result in which an ergodic distribution is similar to the estimate of the density function is most likely obtained when the movement after distribution moves from a higher income in the first five years to lower income gradually.

Table 2 is the result of a similar calculation when the grid line is assumed from the information on the density function. Corresponding grid lines are -0.625, -0.125 and 0.325 in this case and the mean value is included in the third lowest income state. An original number of samples are 365, 451, 255, 129, respectively, and it is found that the sample of the highest income state is negligible.[16] As a result, it is found that there is roughly no major difference when calculating from the estimation results. Ergodic distribution covers 0.2045, 0.2260, 0.2390, and 0.3306 during 1990–2005. The highest income state has the highest density which is still lower than the same result in Table 1. Except for the ergodic distribution in 1995–2000 which differs greatly from the same result of Table 1, the ergodic distribution is gradual in each panel.

[16] It means that the population weight is large in the highest state.

Next, we examined the point at which each element of the transition probability matrix was large. Table 3 and Table 4 show the ratio in three regions when the probability of each element is assumed to be one. For example, when the transition probability from the lowest income state to the lowest income state (in a word, the same state) in 1990–1995 is 0.8913 in Table 1, Table 3 shows that this probability is composed of 83% of Jiangsu's sample, 0% of Shanghai's sample, and 17% of Zhejiang's sample. It is found that the income state was changeless in Shanghai because Shanghai is recorded only in the part of the probability thoroughly staying from the highest state as long as this is seen for the period. Therefore, the transition of the Jiangsu and Zhejiang samples are noteworthy.

The transition between a poor state and a rich state is minimal while the state transition within a poor state and within a rich state as seen in Jiangsu is interesting. As for the transition of the second income state and the third income state after 1995, especially, the entire transition consists of the sample of Zhejiang only (see the fourth panel below). Jiangsu is clearly bi-modal in character. On the other hand, there are some transitions between the second and third state in Jiangsu in Table 4. As a result, it seems that the entire ergodic distribution has leveled off somewhat.

6. Concluding Remarks

This article uses the method of distribution dynamics within the metropolitan area of the Yangtze River Delta (YRD) in 75 districts and counties of which the differences in per capita GDP were studied. The main conclusions of the study are as follows:

First, regional income disparity has always existed in the YRD. The disparity was reduced once during the 8th Five Year Plan then it rose again after that. However, it was reduced again somewhat after 2003. Jiangsu's contribution to disparity is the most and growing and the disparity of each sample in Jiangsu province has expanded.

Second, the income distribution structure has changed greatly between 1990 and 1995. It has changed from a bi-modal to a mono-modal (positively skewed) pattern. Income distribution density of the YRD is not bi-modal as is the whole of China's, but distribution is expanding gradually. The internal structure is complex though there is seemingly a mono-modal pattern distribution in the entire region. Especially, Jiangsu shows a bi-modal distribution and the disparity is large. This result also shows that in the YRD, the various cities and counties are relatively independent in their pace of development.

Third, estimated ergodic distribution shows the income distribution is concentrated in the highest income state. The distribution of the mono-modal pattern is shown in the ergodic distribution as well as in the estimation of the density function. Of course, it is different from provincial distribution in China. And the transition between a poor state and a rich state is very small while the state transition within a poor state and within a rich state is evident in Jiangsu. Therefore, there is relative independence in regional development, but long-term the YRD shows the weaker trend of convergence.

References

Abramovitz, M., 1956. "Resource and Output Trends in the US since 1870," *American Economic Review*, 46, pp.5–23.

Aziz, J., and Duenwald, C., 2001. "China's Provincial Growth Dynamics," IMF Working Paper, WP/01/3. January (http://www.imf.org/external/pubs/ft/wp/2001/wp0103.pdf).

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

Baumol, W.J., 1986. "Productivity Growth Convergence and Welfare: What the Long Run Data Show," *American Economic Review*, 76, pp.1072–1085.

Dowrick, S., and Nguyen, D.H., 1989. "OECD Comparative Economic Growth 1950-1985: Catch-up and Convergence", *American Economic Review*, 79, pp.1010-30.

Islam, N., 1995. "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, pp.1127–1170.

Lee K, Pesaran, M.H., and Smith, R., 1997. "Growth and Convergence: A Multicountry Empirical Analysis of the Solow Growth Model," *Journal of Applied Econometrics*, 12, pp.357–392.

Mankiw, N.G., Romer, D., and Weil, D., 1992. "Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, pp.407–438.

Quah, D., 1993. "Empirical Cross-Section Dynamics in Economic Growth," *European Economic Review*, 37, pp.426–434.

Quah, D., 1996a. "Empirics for Economic Growth and Convergence," *European Economic Review*, 40(6), pp.1353–1375.

Quah, D., 1996b. "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics," *Economic Journal*, 106, pp.1045–1055.

Quah, D., 1996c. "Convergence Empirics across Economies with (Some) Capital Mobility," *Journal of Economic Growth*, 1, pp.95-124.

Quah, D., 1997. "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs," *Journal of Economic Growth*, 2, pp.27–59.

Raiser, M., 1998. "Subsidizing Inequality: Economic Reforms, Fiscal Transfer and Convergence Across Chinese Provinces," *Journal of Development Studies*, 34(3), pp.1–26.

Sakamoto, H., and Islam, N., 2008, "Convergence across Chinese Provinces: An Analysis Using Markov Transition Matrix," *China Economic Review*, 19(1), pp.66–79.

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

Wolff, E.N., 1991. "Capital Formation and Productivity Convergence," *American Economic Review*, 81, pp.565–579.

Xu, Xianxiang, and Li, Xun, 2004. "The Chinese city's convergence of economic growth," *Economic Studies*, 5 (in Chinese).

Figure 1 Map of the YRD (City Level)

Figure 2 Decomposition of Disparity (Mean Log Deviation)

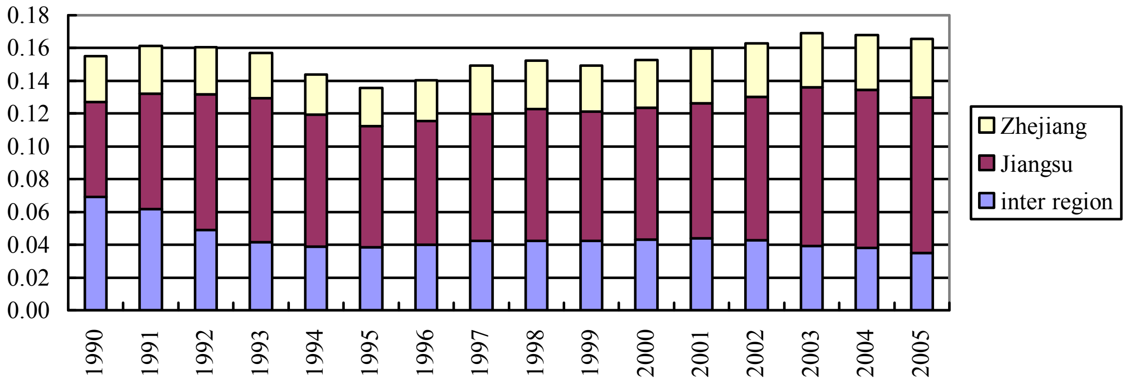

Figure 3 Income Distribution Density in Yangtze River Delta (1990, 1995, 2000, 2005)

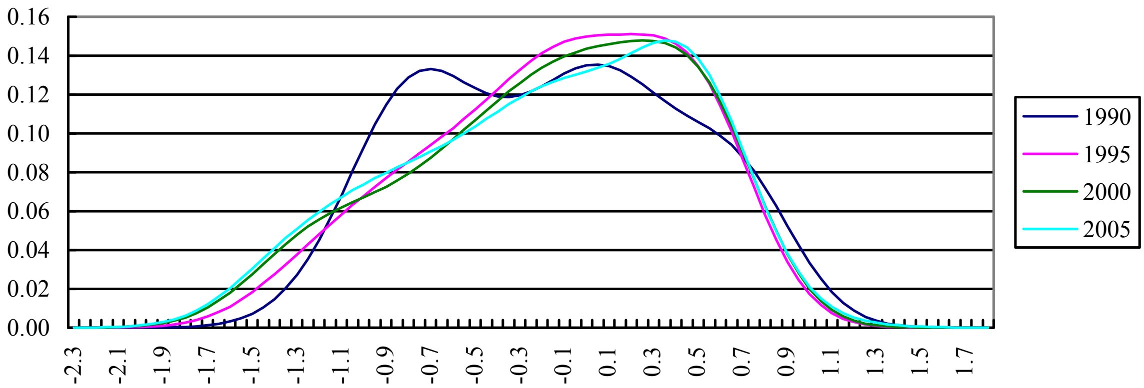

Figure 4 Income Distribution Density in 1990 (Decomposition of 3 Regions)

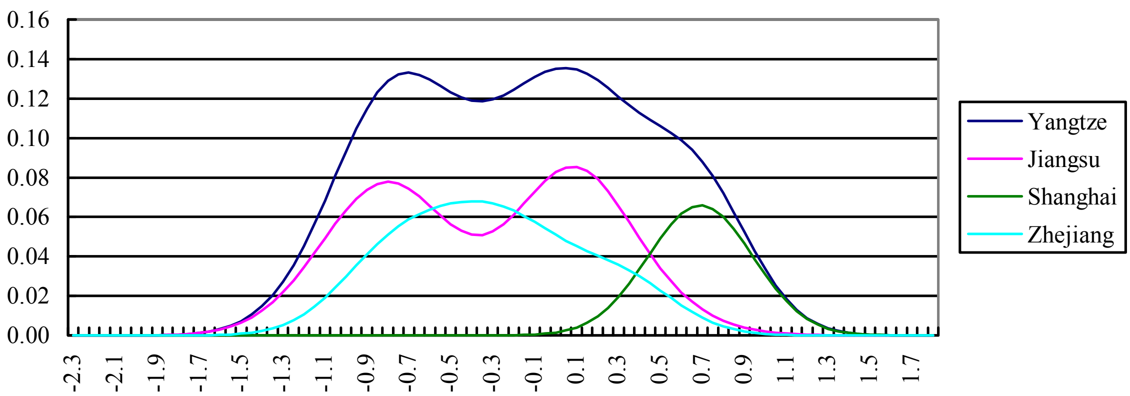

Figure 5 Income Distribution Density in 1995 (Decomposition of 3 Regions)

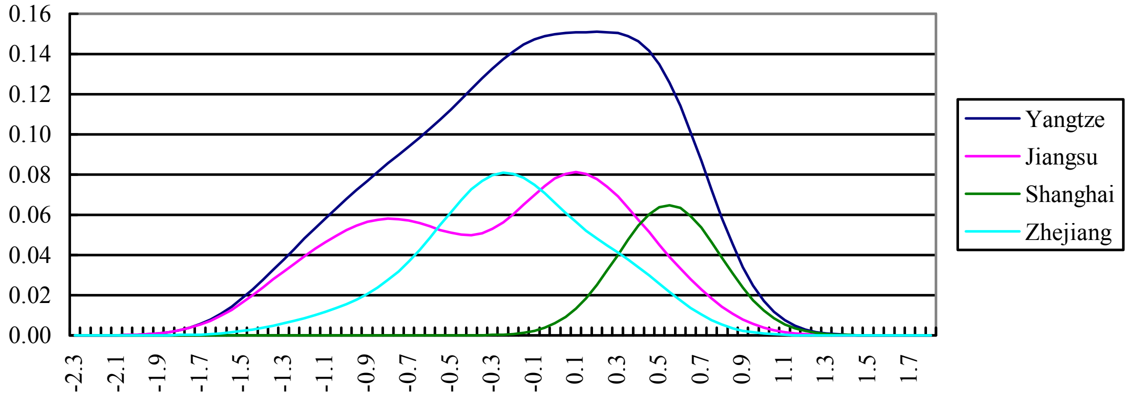

Figure 6 Income Distribution Density in 2000 (Decomposition of 3 Regions)

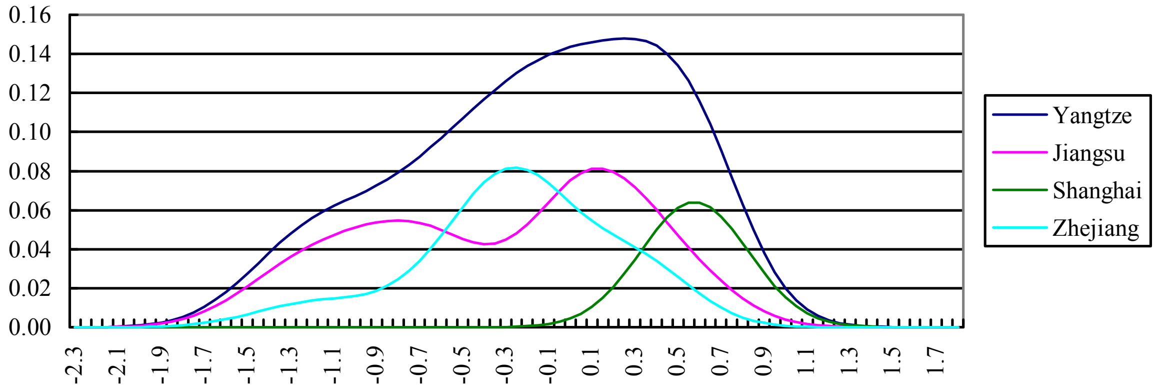

Figure 7 Income Distribution Density in 2005 (Decomposition of 3 Regions)

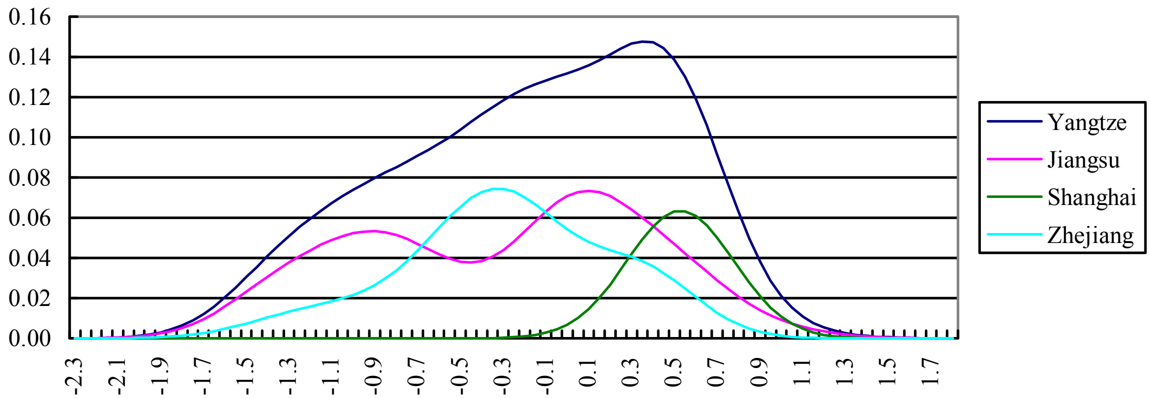

Table 1 Markov Transition Matrix and its Ergodic Distribution (from data information)

		Grid line (highest point)			
		-0.710	-0.337	-0.033	∞
1990-1995	-0.710	0.8913	0.1087	0.0000	0.0000
	-0.337	0.0645	0.7433	0.1922	0.0000
	-0.033	0.0000	0.1107	0.7777	0.1116
	∞	0.0000	0.0000	0.0244	0.9756
	Ergodic	0.0526	0.0887	0.1541	0.7046
1990-2000	-0.710	0.9328	0.0672	0.0000	0.0000
	-0.337	0.0446	0.8382	0.1172	0.0000
	-0.033	0.0000	0.0739	0.8522	0.0739
	∞	0.0000	0.0000	0.0224	0.9776
	Ergodic	0.0782	0.1179	0.1870	0.6169
1990-2005	-0.710	0.9512	0.0488	0.0000	0.0000
	-0.337	0.0554	0.8558	0.0888	0.0000
	-0.033	0.0000	0.0763	0.8622	0.0615
	∞	0.0000	0.0000	0.0228	0.9772
	Ergodic	0.1763	0.1553	0.1808	0.4876
1995-2000	-0.710	0.9823	0.0177	0.0000	0.0000
	-0.337	0.0222	0.9452	0.0326	0.0000
	-0.033	0.0000	0.0422	0.9163	0.0415
	∞	0.0000	0.0000	0.0205	0.9795
	Ergodic	0.2732	0.2178	0.1683	0.3407
1995-2005	-0.710	0.9862	0.0138	0.0000	0.0000
	-0.337	0.0505	0.9160	0.0335	0.0000
	-0.033	0.0000	0.0614	0.8988	0.0398
	∞	0.0000	0.0000	0.0221	0.9779
	Ergodic	0.5914	0.1616	0.0882	0.1588
2000-2005	-0.710	0.9899	0.0101	0.0000	0.0000
	-0.337	0.0760	0.8897	0.0343	0.0000
	-0.033	0.0000	0.0808	0.8812	0.0380
	∞	0.0000	0.0000	0.0237	0.9763
	Ergodic	0.7814	0.1038	0.0441	0.0707

Table 2 Markov Transition Matrix and its Ergodic Distribution (from density information)

		Grid line (highest point)			
		-0.625	-0.125	0.325	∞
1990-1995	-0.625	0.9099	0.0901	0.0000	0.0000
	-0.125	0.0199	0.8635	0.1166	0.0000
	0.325	0.0000	0.0766	0.8303	0.0931
	∞	0.0000	0.0000	0.0491	0.9509
	Ergodic	0.0392	0.1776	0.2704	0.5127
1990-2000	-0.625	0.9366	0.0634	0.0000	0.0000
	-0.125	0.0398	0.8965	0.0637	0.0000
	0.325	0.0000	0.0512	0.9007	0.0481
	∞	0.0000	0.0000	0.0362	0.9638
	Ergodic	0.1387	0.2210	0.2750	0.3653
1990-2005	-0.625	0.9484	0.0516	0.0000	0.0000
	-0.125	0.0467	0.9093	0.0440	0.0000
	0.325	0.0000	0.0416	0.9158	0.0426
	∞	0.0000	0.0000	0.0308	0.9692
	Ergodic	0.2045	0.2260	0.2390	0.3306
1995-2000	-0.625	0.9705	0.0295	0.0000	0.0000
	-0.125	0.0571	0.9251	0.0178	0.0000
	0.325	0.0000	0.0295	0.9608	0.0097
	∞	0.0000	0.0000	0.0226	0.9774
	Ergodic	0.5096	0.2633	0.1589	0.0682
1995-2005	-0.625	0.9713	0.0287	0.0000	0.0000
	-0.125	0.0587	0.9299	0.0114	0.0000
	0.325	0.0000	0.0260	0.9539	0.0201
	∞	0.0000	0.0000	0.0213	0.9787
	Ergodic	0.5248	0.2566	0.1125	0.1062
2000-2005	-0.625	0.9719	0.0281	0.0000	0.0000
	-0.125	0.0604	0.9350	0.0046	0.0000
	0.325	0.0000	0.0221	0.9463	0.0316
	∞	0.0000	0.0000	0.0201	0.9799
	Ergodic	0.5833	0.2714	0.0565	0.0888

Table 3 Decomposition of Probability (from data information)

		-0.710			-0.337			-0.033			∞		
		JS	SH	ZJ	JS	SH	ZJ	JS	SH	ZJ	JS	SH	ZJ
1990-1995	-0.710	0.83	0.00	0.17	0.47	0.00	0.53	0.00	0.00	0.00	0.00	0.00	0.00
	-0.337	0.75	0.00	0.25	0.27	0.00	0.73	0.26	0.00	0.74	0.00	0.00	0.00
	-0.033	0.00	0.00	0.00	0.38	0.00	0.62	0.26	0.00	0.74	0.56	0.00	0.44
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.53	0.00	0.47	0.47	0.37	0.17
1990-2000	-0.710	0.82	0.00	0.18	0.51	0.00	0.49	0.00	0.00	0.00	0.00	0.00	0.00
	-0.337	0.77	0.00	0.23	0.40	0.00	0.60	0.22	0.00	0.78	0.00	0.00	0.00
	-0.033	0.00	0.00	0.00	0.26	0.00	0.74	0.19	0.00	0.81	0.61	0.00	0.39
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.61	0.00	0.39	0.47	0.35	0.18
1990-2005	-0.710	0.81	0.00	0.19	0.55	0.00	0.45	0.00	0.00	0.00	0.00	0.00	0.00
	-0.337	0.71	0.00	0.29	0.38	0.00	0.62	0.19	0.00	0.81	0.00	0.00	0.00
	-0.033	0.00	0.00	0.00	0.17	0.00	0.83	0.22	0.00	0.78	0.63	0.00	0.37
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.00	0.32	0.47	0.35	0.18
1995-2000	-0.710	0.81	0.00	0.19	0.85	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00
	-0.337	0.83	0.00	0.17	0.51	0.00	0.49	0.00	0.00	1.00	0.00	0.00	0.00
	-0.033	0.00	0.00	0.00	0.00	0.00	1.00	0.14	0.00	0.86	0.74	0.00	0.26
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.00	0.30	0.48	0.34	0.19
1995-2005	-0.710	0.79	0.00	0.21	0.90	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00
	-0.337	0.68	0.00	0.32	0.43	0.00	0.57	0.00	0.00	1.00	0.00	0.00	0.00
	-0.033	0.00	0.00	0.00	0.00	0.00	1.00	0.20	0.00	0.80	0.71	0.00	0.29
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.77	0.00	0.23	0.47	0.34	0.18
2000-2005	-0.710	0.78	0.00	0.22	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	-0.337	0.64	0.00	0.36	0.36	0.00	0.64	0.00	0.00	1.00	0.00	0.00	0.00
	-0.033	0.00	0.00	0.00	0.00	0.00	1.00	0.27	0.00	0.73	0.69	0.00	0.31
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.83	0.00	0.17	0.47	0.35	0.18

Table 4 Decomposition of Probability (from density information)

		-0.625			-0.125			0.325			∞		
		JS	SH	ZJ	JS	SH	ZJ	JS	SH	ZJ	JS	SH	ZJ
1990-1995	-0.625	0.78	0.00	0.22	0.37	0.00	0.63	0.00	0.00	0.00	0.00	0.00	0.00
	-0.125	0.00	0.00	1.00	0.26	0.00	0.74	0.33	0.00	0.67	0.00	0.00	0.00
	0.325	0.00	0.00	0.00	0.36	0.00	0.64	0.81	0.00	0.19	0.60	0.00	0.40
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.00	0.65	0.20	0.63	0.17
1990-2000	-0.625	0.79	0.00	0.21	0.46	0.00	0.54	0.00	0.00	0.00	0.00	0.00	0.00
	-0.125	0.59	0.00	0.41	0.22	0.00	0.78	0.35	0.00	0.65	0.00	0.00	0.00
	0.325	0.00	0.00	0.00	0.36	0.00	0.64	0.75	0.00	0.25	0.64	0.00	0.36
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.51	0.00	0.49	0.23	0.63	0.14
1990-2005	-0.625	0.78	0.00	0.22	0.50	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00
	-0.125	0.55	0.00	0.45	0.21	0.00	0.79	0.33	0.00	0.67	0.00	0.00	0.00
	0.325	0.00	0.00	0.00	0.36	0.00	0.64	0.72	0.00	0.28	0.73	0.00	0.27
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.61	0.00	0.39	0.25	0.63	0.12
1995-2000	-0.625	0.80	0.00	0.20	0.77	0.00	0.23	0.00	0.00	0.00	0.00	0.00	0.00
	-0.125	0.77	0.00	0.23	0.20	0.00	0.80	0.46	0.00	0.54	0.00	0.00	0.00
	0.325	0.00	0.00	0.00	0.34	0.00	0.66	0.71	0.00	0.29	1.00	0.00	0.00
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.86	0.00	0.14	0.26	0.64	0.10
1995-2005	-0.625	0.79	0.00	0.21	0.74	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.00
	-0.125	0.63	0.00	0.37	0.19	0.00	0.81	0.37	0.00	0.63	0.00	0.00	0.00
	0.325	0.00	0.00	0.00	0.35	0.00	0.65	0.69	0.00	0.31	1.00	0.00	0.00
	∞	0.00	0.00	0.00	0.00	0.00	0.00	0.93	0.00	0.07	0.28	0.62	0.10
2000-2005	-0.625	0.77	0.00	0.23	0.72	0.00	0.28	0.00	0.00	0.00	0.00	0.00	0.00
	-0.125	0.50	0.00	0.50	0.18	0.00	0.82	0.00	0.00	1.00	0.00	0.00	0.00
	0.325	0.00	0.00	0.00	0.36	0.00	0.64	0.67	0.00	0.33	1.00	0.00	0.00
	∞	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.30	0.61	0.10