Weighted convergence and regional clusters across China*

M. J. Herrerías† **Vicente Orts**† **Emili Tortosa-Ausina**‡

December 11, 2009

Abstract

We analyse per capita income convergence among 28 Chinese provinces between 1952 and 2005 using the distribution dynamics approach. Compared with previous studies, we provide a more complete view by including some additional information such as the asymptotic half life of convergence, mobility indices and the continuous version of the ergodic distributions. In addition, we also extend the analysis to evaluate whether patterns could differ if weighted by either the population living in each province or their economic sizes, together with the existence and magnitude of spatial spillovers. The unweighted, unconditional analysis corroborates and supplements previous findings, especially those indicating that convergence patterns differ strongly under either pre- or post-reform trends. Both the weighted and space-conditioned analyses indicate that convergence could be much faster when these factors are introduced in the analysis. Implications are especially relevant when weighting by population, since results indicate that the number of people escaping from relative poverty would be much higher than the figure predicted by the unweighted analysis.

Key words and phrases: China, convergence, distribution dynamics, provinces, weights

JEL Classification: C16, O18, O47, R11

Communications to: M.J. Herrerías, Departament d'Economia, Universitat Jaume I, Campus del Riu Sec, 12071 Castelló, Spain. Tel.: +34 964388615, fax: +34 964728591, e-mail: `herreria@uji.es`

*M.J. Herrerías and Vicente Orts gratefully acknowledge the financial support of the Ministerio de Ciencia e Innovación (ECO2008-06057/ECON) and Generalitat Valenciana. Emili Tortosa-Ausina acknowledges the financial support of Fundació Caixa Castelló-Bancaixa (P1.1B2008-46), Ministerio de Ciencia e Innovación (ECO2008-03813/ECON and ECO2008-05908-C02-01/ECON) and Generalitat Valenciana (PROMETEO/2009/066). The usual disclaimer applies.

†Universitat Jaume I.

‡Universitat Jaume I and Ivie.

1. Introduction

Income disparity across Chinese regions is a major concern among policymakers, and also an interesting case of study from an academic point of view. Growing inequality increases redistributive tax pressures, which deters investment incentives and can also lead to a more unstable socio-political environment for economic activities (Alesina and Perotti, 1993; Alesina and Rodrik, 1994). Given its potential to thwart both economic growth and stabilisation, inequality and poverty reduction across regions is one of the fundamental problems that the Chinese government must solve and, accordingly, several initiatives have been undertaken to promote income distribution across Chinese regions. One of the most prominent ones took place in the early 2000s, when the Chinese government launched the *Great Western Development Program* (GWDP) with the aim of investing more in the western regions, where economic development is lower. The purpose of this programme was to balance the degree of development across regions and to reduce poverty, but this expectation has not been fulfilled as of today. Therefore, examining whether convergence is taking place across Chinese provinces is not only of great significance because of the sheer number of people whose welfare is involved (China represents about one-fifth of the world's population), but also because it makes it possible to evaluate the success of policies designed to alleviate the magnitude of the inequalities. In addition to this, China has an economy that is undergoing a transition from a centrally planned to a more market-oriented economy, and which has its own particular characteristics. Therefore, as indicated by Sakamoto and Islam (2008), the Chinese case could help to assess whether switching from central planning to a market mechanism makes a difference with regard to convergence. This could be increased further to examine whether interregional differences in income levels tend to disappear or to increase over time as a result of this transformation.

As indicated by Islam (2003), there are different definitions of convergence, each of which is linked to growth theory in a different way. According to this author, not only are there different ways in which convergence can be understood but also different methodologies to evaluate it. The definitions would include convergence *within* an economy vs.*across* economies, *growth rates* versus income *level* convergence, β-convergence versus σ-convergence, unconditional versus conditional convergence, global versus local or club-convergence, and deterministic versus stochastic convergence. The methods would be the informal cross-section approach, the formal cross-section approach, the panel approach, the time-series approach, and the distribution approach. This is perhaps the most comprehensive and up-to-date survey, but given the importance of the issue, other significant works have also been published such as Quah (1997b), De la Fuente (1997) Durlauf and Quah (1999), or Temple (1999).

From these surveys it can be seen that there is a substantial body of theoretical and empirical research focusing on the issue of country and regional convergence, and the Chinese case is by no means an exception. Most previous studies have examined convergence across Chinese provinces using parametric techniques which adopt either cross-section or panel data approaches (see, for example Rozelle, 1994; Jian et al., 1996; Chen and Fleisher, 1996; Raiser, 1998; Yao and Zhang, 2001a,b; Weeks and Yao, 2003;

Wang, 2003; Pedroni and Yao, 2006).[1] However, the empirical evidence found in the literature on the subject is rather mixed. For example, Rozelle (1994) found divergence within the Jiangsu province for the 1984–1989 period and (Jian et al., 1996, p. 8) concluded that during the pre-reform period (prior to 1978) the Chinese provinces tended to diverge, while in the post-reform period (after 1978) until 1990, a tendency towards absolute convergence predominated. Nevertheless, this pattern changes to divergence in per capita incomes from 1990 to 1993. Using cross-section and panel data techniques, Chen and Fleisher (1996) provide evidence of divergence in the pre-reform period and convergence from 1978 to 1993.[2] In contrast, opposite results were found by Weeks and Yao (2003) using the GMM estimator. In addition, despite using different methodological approaches, Yao and Zhang (2001a,b) found that the Chinese regions did not converge in the reform period (1978–1995). Indeed, their results clearly indicated evidence of divergence in the different geo-economic clubs (coastal and non-coastal zones). Only when they controlled for regional effects and other determinants of growth, did they find conditional β-convergence using the panel data approach (Yao and Zhang, 2001b). More recently, in a study covering the 1952–1997 period, and using non-stationary panel techniques, Pedroni and Yao (2006) provided empirical support for the fact that the long-run tendency since the reforms has been for provincial-level incomes to continue to diverge. They added that this divergence cannot be attributed to the presence of separate, regional convergence clubs divided among common geographic sub-groupings such as the coastal versus interior provinces or preferential policies. This ambiguity in the results, of course, depends on the selection of the period under study, the estimation method that is used, the variables that the researcher has considered, and in studies that examine the differences in per capita income among clubs, the *a priori* selection of these small groups of regions, could affect the empirical analysis (Maasoumi and Wang, 2008).

In contrast, there is little empirical evidence of convergence across the Chinese provinces using the distribution dynamics model developed by Quah (1993a). Only two recent papers apply this new way of analysing economic convergence across the Chinese provinces: Bhalla et al. (2003) and Sakamoto and Islam (2008).The former investigated convergence patterns from 1952 to 1997 using per capita data among Chinese provinces. They concluded that there is evidence of convergence within the pre-defined geo-economic sub-regions, but no evidence was found of convergence between the sub-regions. In particular, they argued that the gap between the eastern and the central regions was small in the pre-reform period, but widened rapidly in the reform period. The same pattern occurs with eastern and western provinces, but with a more significant fluctuation over time. These results imply a strong divergence between these two pairs of regions. More recently, the latter authors—Sakamoto and Islam (2008)—found similar results. Indeed, their findings indicated that the distribution of per capita income across Chinese provinces over time has attained a bimodal characteristic with two opposing tendencies in the two sub-periods considered (1952–1978 and 1978–2003). During the pre-reform period the dynamics of the distribution indicated that there were more provinces piling at lower values of per capita income, whereas during the post-reform

[1] Studies such as Rozelle (1994), Jian et al. (1996), Gundlach (1997), and Yao and Zhang (2001a,b) did not consider the endogeneity or the dynamics of the models, and were thus in line with the critique stated by Caselli et al. (1996). Only few works try to deal with this problem, such as Weeks and Yao (2003), Wang (2003) and Ding et al. (2008) among others.

[2] Similar results were found by Raiser (1998) and Wang (2003) for the post-reform period.

period the dynamics of the distribution moved in the opposite direction, namely, there were more provinces moving towards higher income groups. In spite of these results, Sakamoto and Islam (2008) argued that the distribution dynamics of the reform period do not seem to have led to a stable pattern yet, thus making prediction difficult, and hence it remains an open issue to be analysed further.

This paper examines the complexity of the convergence process in per capita income across the 28 Chinese provinces over the period 1952–2005, which means that our proposal therefore stands with those using the distribution approach developed by Quah (1993a,b). Unlike previous studies that apply either σ- or β-convergence in cross section or panel data techniques (which sometimes require strong assumptions) we allow data to reveal the nature of the relationship of interest by using nonparametric techniques. It is mainly a data driven approach, and we do not impose any assumption or restriction on the specification of the density of the distribution. So, in its initial steps, our investigation differs only slightly from that conducted by Sakamoto and Islam (2008).

However, we introduce a series of variations with respect to both Bhalla et al.'s (2003) and Sakamoto and Islam's (2008) proposals. Some of the differences we introduce have to do with the fact that the analyses of Chinese income distribution and convergence have dealt with the behaviour of incomes in terms of *provinces*—i.e., regional convergence. However, as indicated by Jones (1997), while this is a common way to view and analyse data, it can be highly misleading: should provincial borders be drawn differently, conclusions might vary remarkably. Alternatively, we could weight each province by its population (although other weighting schemes are possible) so that the unit of observation was then a *person* instead of a province. As indicated by Sala-i-Martin (2006), the unweighted approach is not useful if one is concerned about human welfare, since different provinces have varying population sizes. In this regard, the most important fact to note is that, for instance, by 2005 the population living in Sichuan was more than 20 times larger than the population living in Qinghai or in Ningxia. Disparities were even higher by 1952—the population living in Sichuan was more than 40 times larger than that living in Ningxia. Therefore, the experience of the most populated provinces largely determines what happens to the "average" person in China.

By weighting by population, some researchers have drawn different conclusions to those reached via unweighted analyses. For instance, Jones (1997) showed that the emergence of a bimodal distribution disappeared once each country data point was weighted by population, whereas Schultz (1998) found that, when one uses population-weights, it is no longer true that incomes tend to diverge. Given the disparities in terms of both population and GDP across Chinese provinces, one may expect some interesting conclusions also to emerge in the case of China when comparing our results to the unweighted analysis by Sakamoto and Islam (2008).

The distribution analysis approach is also attractive because of its ability to disentangle the existence of spatial spillover effects, in a similar fashion to Quah (1996c). Following this author's approach, we measure whether these spillovers could exist or not by evaluating the magnitude of the contiguity effect across Chinese provinces. The rationale for this lies in the fact that the economic development of a particular region could be strongly related to that of its neighbouring provinces. The issue is particularly relevant in the case of China, whose government considered that by developing the coastal regions, the central and western provinces would also boost their development via (spatial) spillover effects. However, empirical evidence evaluating these policies is still scarce. Only Brun et al. (2002) have conducted research on the issue, their findings indicating a relative failure of the growth of the coastal regions from 1981 to 1998 to trigger development in the western provinces. Therefore, according to these authors, it would be wrong to expect spillover effects to be enough to reduce disparities between Chinese provinces, at least in the short run.

The rest of the paper is organised as follows. Section 2 describes the main trends in provincial distribution of per capita GDP in the Chinese economy during the period under consideration. Section 3 deals with the technical aspects of the distribution analysis model. In Section 4 we analyse the results of applying the model to per capita income data for 28 Chinese provinces. Finally, we present some concluding remarks in Section 5.

2. Emerging patterns in provincial distribution of Chinese per capita GDP: 1952–2005

A comprehensive description of the evolution of the Chinese economy is beyond the scope of this section and the reader is directed to other studies, such as Lardy (1992), Chai (1998) or Bramall (2000), for more details. Here, we briefly summarise the most important trends of the per capita GDP and population across Chinese provinces during the period under consideration (1952–2005). Table 1 shows the per capita GDP and population of the 28 Chinese provinces in 1952, 1978 and 2005, the growth rates of both magnitudes for the whole period (1952–2005) and for the two sub-periods considered (1952–1978 and 1978–2005), as well as the corresponding standard deviation of all magnitudes reported. As can be seen, there are substantial differences in per capita GDP among provinces in each year, as well as for each province between 1952 and 2005. In 1952 the average provincial per capita GDP in China was 140.86 Yuan and its standard deviation was 78.07 (coefficient of variation of 0.55). Between 1952 and 2005, the per capita GDP of Chinese provinces, measured in constant 1952 prices, was growing at a cumulative average growth rate of 6.39%, the result being that the average provincial per capita GDP reached 4470.46 Yuan in 2005. However, by that year the coefficient of variation had increased to 1. This rapid growth and the increase in regional disparities have two very different steps in time: prior to the economic reforms, i.e. before 1978, and the post-reform period.

The pre-reform period (1952–1978) was characterised by the central planning of the allocation of economic resources and an unstable political environment. China experienced many booms and boosts, like the Great Leap Forward (1958–1961) or the famine caused by failures in the agricultural sector following the unstable economic and political environment that accompanied the Cultural Revolution (1966–1976). Nevertheless, in spite of these turbulences the average provincial per capita GDP grew at a rate of nearly 4%, although regional inequalities increased significantly in that period. During those years the coefficient of variation of provincial per capita GDP increased from the aforementioned 0.55 to 1.19. The existence of barriers across regions (Rozelle, 1994) probably accounts for the divergence rather than convergence in per capita GDP that took place in the pre-reform period. These barriers are to be understood in the broad sense of the term and are notably related with the mobility of workers, the unequal specialisation in the different economic activities of provinces, the promotion of investment in heavy industry rather than in agriculture or the centralised fiscal system (Wei, 1996).

In contrast, in the post-reform period (1978–2005), the average growth rate of the per capital GDP of the Chinese provinces increased to 8.79%, while the provincial inequality in per capita GDP, as measured by the coefficient of variation, declined from 1.19 to 1. This period is characterised by the economic reforms initiated in the late 1970s, including the progressive adoption of market-oriented and open-door strategies for development (that culminated in 2001 with its adhesion to the WTO), and which were gradually transforming the Chinese economy towards a more market-oriented, decentralised and open economy.

To sum up, these stylised facts reflected the most widely extended result in empirical studies, i.e. the absence of convergence or even divergence during the pre-reform period and slight convergence in the post-reform period.[3]

However, a closer look at the data reveals that, even within each sub-period, the dynamics of the first moments of the distribution of per capita GDP were very volatile, especially the variance throughout the first sub-period, but also in the second sub-period.[4] Although in the period 1978–2005, the variability of the variance of the distribution was much smaller, some changes in its trend can also be seen. Between 1978 and the late 1980s and early 1990s, the variance in distribution drops regularly, then the trend changes and increases again almost to the end of the sample (2004). At the end of the sample a new change in the trend of the variance appears indicating a new decrease in the inter-provincial inequality. Table 1 shows that the performances of the provinces in each economic zone (east, central and west) were also very different to each other throughout the period under consideration. Thus, while the provinces in the eastern zone reproduce the aggregate changes in average growth and inequality of per capita GDP on a different scale, in the central and western zones the dynamics varies markedly between them as well as between the two sub-periods. In the pre-reform period the average growth rate of the western provinces was superior to that of the provinces in the central zone (3.87% and 2.66% respectively), this difference being reversed in the post-reform period (7.79% and 8.77%). Furthermore, the dispersion of the provincial growth rates slowed down significantly in the post-reform period in all the zones. At the same time, while the inequality among provinces in the central zone declined throughout the whole of the period considered (the coefficient of variation was 0.42, 0.37 and 0.31 in 1952, 1978 and 2005 respectively), in the western provinces the average inequalities did not change in the pre-reform period and increased significantly in the post-reform period (the coefficients of variation were 0.30, 0.30 and 0.49 respectively). The picture is

[3] See, for example, Bhalla et al. (2003).

[4] We do not report the standard deviation of the all distributions year by year, but they are available upon request.

complex, and when we look at the data in more detail, the more differences we find in the performance of different provinces and zones. In fact, as stated by Quah (1996b,c,d, 1997a), looking only at the first and second moments of the distribution is likely to be uninformative in the case of multimodal distributions, as could be the case, and therefore it is better to analyse the entire distribution of provincial per capita GDP and its dynamics.

3. Per capita income convergence as distribution dynamics

The literature on growth and convergence has debated intensely about the importance of analysing distribution dynamics in order to understand the mechanics of economic development. This view was strongly supported not only by (Quah, 1993a, 1996b,d) but also by many others who advocate the analysis of the dynamics of the entire cross-section distribution of per capita income (or labour productivity). The reasons for this lie in the fact that uncovering *all* the information on the dynamics using only summary statistics is a questionable procedure. Accordingly, empirical studies have shown consistent evidence of a cross-country income distribution displaying bimodality with a marked thinning in the middle. This transformation implied that while by the 1960s many countries belonged to the middle income group, by the 1990s the world polarised into two groups, namely the rich and the poor, a phenomenon to which Quah (1996d) refers to as "twin-peak" or "two-club" convergence. However, as indicated by Cetorelli (2002), there is a positive probability of an economy moving from one group to the other, i.e. the bimodal distribution is ergodic. One can therefore observe previously poor economies that grow rapidly and move to join the rich club; reversals of fortune, where fast growth is only temporary and may be followed by abrupt halts and decumulation; or economic disasters involving previously rich economies regressing to lower levels of income (Cetorelli, 2002).

The instruments provided by Quah (1993a,b), along with some others borrowed from the literature on income inequality (Shorrocks, 1978), are of remarkable interest for analysing provincial per capita income dynamics in China. Quah's critique to previous approaches to examining convergence (basically those based on analysing β- and σ-convergence) points out that conclusions are based on (two) summary statistics only. However, both the mean and the standard deviation give an interesting but incomplete illustration of the entire distribution of per capita income, for it conceals some significant features such as the existence of multiple modes. This and related phenomena would be overlooked unless an analysis taking into account different groups of provinces were performed; however, focusing on the *entire* distribution is even better than carrying out the analysis for different groups of provinces.

3.1. Intra-distribution mobility and ergodic distributions

Our variable of interest is the normalised logarithm of per capita income, i.e. divided by the mean for the 28 provinces.[5]

[5]For the sake of simplicity, we use the concept of provinces throughout this paper. However, in China there are 23 provinces, 5 autonomous regions, 4 municipalities, and 2 special administration regions (SAR). We have excluded Tibet

We consider this type of normalisation because of the informativeness of its interpretation: the closer a value is to the unity, the closer it will be to the national average. Therefore, the more values there are close to unity, the higher the convergence to this national average will be. Our selected variable is the same as the one chosen by Sakamoto and Islam (2008), but it is normalised in a slightly different way—they use the log of normalised (divided by the mean) per capita income. While it is basically the same variable, we consider that ours has the interesting feature of being more directly interpretable—we can measure, for instance, whether a particular province has twice or half the average. While the normalisation selected by Sakamoto and Islam (2008) provides similar information, it is not as direct. Like Sakamoto and Islam (2008), we denote this variable by x_i, so that $x_i = \log y_i / \log \bar{y}$, where y_i is the per capita income of province i, and $\bar{y}$ is the cross-section average of y_i.

Therefore, in our setting, $s_{i,t}$ refers to province i's normalised per capita income in period t, whereas $F_t(s)$ refers to the cumulative distribution of $s_{i,t}$ across provinces. Corresponding to $F_t(s)$ we can define a probability measure λ_t s.t.:

$$\lambda_t((-\infty, s]) = F_t(s), \ \forall s \in \mathbb{R}. \tag{1}$$

where λ_t is the probability density function for each indicator across provinces in period t. Therefore, the model analyses the dynamics of λ_t, i.e., the dynamics of the cross-section distribution of per capita income,[6] for which we consider a stochastic difference equation:

$$\lambda_t = P^*(\lambda_{t-1}, u_t), \ \text{integer } t, \tag{2}$$

In the above equation, $\{u_t : \text{integer } t\}$ is the sequence of disturbances of the entire distribution, and P^* is the operator mapping disturbances and probability measures into probability measures. In other words, the P^* operator reveals information on how the distribution of per capita income at time $t-1$ (y_{t-1}) transforms into a different distribution at time t (y_t).

Following Redding (2002), we may assume that the stochastic difference equation is of first order and that operator P^* is time invariant. Thus, by setting null values to disturbances and iterating in (2) we obtain the future evolution of the distribution:

$$\lambda_{t+\tau} = (P^* \cdot P^* \cdot \ldots \cdot P^*)\lambda_t = (P^*)^\tau \lambda_t \tag{3}$$

By discretising the set of possible values of s into a finite number of cells $k \in \{1, \ldots, K\}$, P^* becomes

due to the lack of data. Moreover, this paper focuses on Mainland China, and consequently we have also excluded Taiwan, Hong Kong and Macao. In addition, Hainan is included within Guangdong province and Chongqing is included as part of Sichuan province, given that the former was separated from Guangdong in 1988, and the latter was part of Sichuan until 1997. This is standard practice in Chinese studies. Related to this, there is a debate in the literature about the quality of Chinese statistics. However, it is possible to find support for the quality of the Chinese statistics required to examine the long-run trends in Holz (2005), Chow (2006) and Bai et al. (2006). We use one of the latest and revised compilations edited by National Bureau of Statistics of China (NBS) in 2005 and 2006, which provides us with information that is homogenous enough, both across Chinese provinces and over time, to perform this study properly. More specifically, our main data source is "China Compendium of Statistics, 1949-2004".

[6] From now on, when talking about per capita income we will be referring implicitly to normalised log of per capita income.

a transition probability matrix

$$\lambda_{t+1} = P^* \cdot \lambda_t \tag{4}$$

In this transition probability matrix, λ_t turns into a $K \times 1$ vector of probabilities that a given province per capita income is located in a given grid at time t.

Discretisation divides the space of possible F_t values into discrete grid cells (what some authors call "states" or "classes") e_k, $k = 1, \ldots, K$. Then, after classifying each country-year observation into one of the K states, we build up a 5×5 matrix whose p_{kl} entries indicate the probability that a country that is initially in state k will transit to state l during the period or periods considered (T). Each row of the matrix constitutes a vector of transition probabilities, which adds up to unity. The boundaries between grid cells are chosen so that country-year observations are divided approximately equally among the cells, each cell corresponding to approximately one fifth of the distribution of the selected variable across provinces and time. Interpretation is straightforward: observations in the first state refer to the poorest provinces. This way of constructing is common practice (see, for instance Redding, 2002; Lamo, 2000). Some other contributors have considered different criteria such as selecting the limits between states arbitrarily—although reasonably (Kremer et al., 2001; Quah, 1993a). Alternatively, it is possible to dodge the discretisation problem by considering stochastic kernels (Quah, 1996c), although these present some difficulties for estimating the ergodic, or stationary distribution.

Transition probability matrices make it possible to measure the probability of a given province moving to a higher (or lower) position on the grid. Calculating the transition probability matrices starts by discretising the set of observations into the selected states e_k. The interpretation of the different figures in each matrix is straightforward. In the case of the limits between states, those for which $e_k = (0.25, 0.50)$ would include provinces whose per capita income ranged between one quarter and half the national average. In the case of the different entries in the matrices, they indicate the probability of a given province transiting out from its initial state to other states during the period or periods considered.

To compute each transition matrix we count the number of transitions out of and into each cell, i.e., for each p_{kl} cell:

$$p_{kl} = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{n_{kl}^t}{n_k^t} \tag{5}$$

where T is the number of years or periods, n_{kl}^t is the number of countries moving during one period from class k to class l, and n_k^t is the total number of provinces starting the period in class l.

According to this methodology, transitions are estimated by counting the number of provinces moving from one class to another. However, as indicated in the introduction, using provinces as units of analysis would not be useful if we were concerned with human welfare because different provinces have different population sizes. Therefore, the unweighted analysis does not help to answer questions such as "How many people in China live in poverty?" or "How have poverty rates changed over the last few decades?" Therefore, it is also relevant to estimate *weighted* transition probability matrices, for which different weighting schemes are feasible, and are not limited to just population. The underlying idea is that the impact on Chinese per capita income will be greater if a larger country transits out than if a small province does so. Therefore, we count provinces' transitions, but in this case each province is represented by its entire share of Chinese population (in the case of population-weighted transition probability matrices), so that the unit of observation is now a *person* instead of a country, i.e. we count the number of persons moving between states. This issue is often ignored, although exceptions do exist, such as Kremer et al. (2001) or Jones (1997).

By operating with the information offered by the transition probability matrix we can also characterise the hypothetical long-term ergodic or stationary distribution. The variety of resulting scenarios might be remarkable, including distributions with the probability mass concentrated mainly in the central classes (indicative of convergence to the mean if these central states contained the unity), polarised distributions ("twin peaks") indicating that the poorest and richest are becoming increasingly more distant from each other, or one with the probability mass distributed in the extreme classes (tails) of the distribution. Therefore, the ergodic distributions make it possible to determine the predominating long-run tendency for provincial per capita income in China.

3.2. Transition path analysis and mobility indices

We can also evaluate the speed with which the ergodic distribution, or steady-state, is approached by means of the concept of asymptotic half life of the chain, $H - L$, which is how long it takes to cover half the distance from the stationary distribution. Following Shorrocks (1978), the half life is defined as:

$$H - L = -\frac{\log 2}{\log |\lambda_2|} \tag{6}$$

where $|\lambda_2|$ is the second largest eigenvalue (after 1) of the transition probability matrix. It ranges between infinity—when the second eigenvalue is equal to 1 and the stationary distribution does not exist—and 0—when $\lambda_2 = 0$ and the system has already reached its stationary equilibrium (Magrini, 1999).

We also consider the mobility indices proposed by the literature on economic inequality (Shorrocks, 1978; Geweke et al., 1986). As suggested by Quah (1996a), analogously to the measures of income inequality designed to collapse the information contained in an entire distribution into a single scalar, a mobility index summarises the mobility information in a transition probability matrix into one number. We consider the proposals by Shorrocks (1978) and Geweke et al. (1986), summarised by Quah (1996a). In their proposals, by defining the mobility index as a continuous real function $\mu(\cdot)$ over the set of transition matrices $\mathcal{P}$, it satisfies the properties of normalization, monotonicity, immobility, and perfect mobility (see Shorrocks, 1978). This index (μ^1) evaluates the trace of the transition probability matrix, disclosing information on the relative magnitude of diagonal and off-diagonal terms. It is identical to the inverse of the harmonic mean of expected durations of remaining in a certain state and, following Quah (1996a), its particular expression is:

$$\mu_1(P^*) = \frac{K - \mathrm{tr}(P^*)}{K-1} = \Big(\frac{K}{K-1}\Big)\Big\{K^{-1}\sum_j (1 - P^*_{jj})\Big\} = \frac{K - \sum_j \lambda_j}{K-1} \tag{7}$$

where K is the number of classes, P^*_{jj} is the j-diagonal entry of matrix P^*, which represents the probability of remaining in state j, and λ_j represent eigenvalues of P^*. Large values of μ_1 indicate less persistence (or more mobility) in P^*.

3.3. The evolution of the external shape of the distributions

It is also relevant to provide information on both the initial and final distributions for the variable of interest, in order to gain further insights into how distributions have evolved. Therefore, for all indicators we provide four sets of additional results, namely, transition probability matrices, ergodic distributions, initial distributions, and final distributions.

However, in their present form, the three sets of distributions share a common disadvantage, namely, they are discrete and probability is spread out across one set of states only. Although we have provided reasons why such a disadvantage may not be as restrictive as some authors suggest, we try to be as informative as possible by also providing the continuous counterpart to this discrete estimation, namely, the non-parametric estimation of density functions via kernel smoothing. This is the first step in Quah's model of distribution dynamics, and it provides remarkable insights about the convergence process. If the probability mass became tighter, it would indicate convergence, whereas if it became flatter, it would be indicative of divergence. As can be easily inferred, multiple scenarios may result.

We consider a kernel estimator for each indicator:

$$\hat{f}(x) = \frac{1}{Nh}\sum_{i=1}^{N} K\Big(\frac{\|x - X_i\|_x}{h}\Big) \tag{8}$$

where x is the point of evaluation, X is the indicator of interest, N is the number of observations (countries), h is the bandwidth, $\|\cdot\|_x$ is a distance metric on the space of X, and $K(x)$ is a kernel function (see Härdle and Linton, 1994) which are generally required to hold that:

$$\int_{\mathbb{R}} K(x)dx = 1, \quad \int_{\mathbb{R}} xK(x)dx = 0, \quad \sigma^2_K = \int_{\mathbb{R}} x^2 K(x)dx < \infty \tag{9}$$

There are several choices for $K(x)$, which may be defined in terms of univariate and unimodal probability density functions. For the sake of simplicity, we consider a Gaussian kernel:

$$K(x) = (1/\sqrt{2\pi})e^{-\frac{1}{2}x^2} \tag{10}$$

Weighting densities (in order to provide continuous counterparts to the weighted initial and final distributions) requires slight modifications. Few studies have considered this, despite its potential relevance in some specific contexts. Following Goerlich (2003), expression (8) is slightly modified to become:

$$\hat{f}_\omega(x) = \frac{1}{h}\sum_{i=1}^{N}\omega_i K\Big(\frac{\|x - X_i\|_x}{h}\Big) \tag{11}$$

where ω_i is the share of either world output or world population (depending on the type of weighting we consider) corresponding to country i.

The continuous version of the ergodic distributions is more difficult to estimate. In this case, related literature is scarce. Some studies provide estimations for ergodic densities (see Johnson, 2000, 2005). However, no studies provide, simultaneously, results for ergodic distributions yielded by transition probability matrices *and* ergodic densities. In order to obtain a fully compatible view between the results of the transition probability matrices and their continuous counterpart, we generated ergodic densities considering the information in the (discretised) ergodic distributions (1×20). Specifically, we generated normal distributions for each of the twenty states over which probability is spread out, with a number of observations proportional to each state's share of ergodic probability. This generates a pseudo-histogram in which we do not have bars, but normal distributions. Then we proceed in exactly the same way as when smoothing both initial and final distributions, i.e. by considering kernel methods to smooth the observations in each of these twenty states. This algorithm yields ergodic densities which are fully consistent with the ergodic distributions computed from transition probability matrices. The continuous state approach naturally complements the view provided by discrete ergodic distributions, which tend to summarise too much information in a few states. Although the information provided by ergodic densities is essentially the same, we remove the arbitrariness implied by selecting a grid.

3.4. Conditioning on neighbor-relative information

The techniques employed enable us to analyse the importance of spatial factors in explaining regional convergence or divergence. In particular, following Quah (1996c), we can analyse the role played by surrounding regions in explaining the dynamics of regional distribution of per capita GDP (conditional distribution dynamics). The specific hypothesis to be tested is whether Chinese provinces might be converging with their neighbours, i.e. the provinces around them. Quah (1996c, 1997a) provides reasons as to why such a convergence pattern could exist, along with methods to evaluate conditional convergence with our instruments. As indicated by Quah (1996c), and most of the literature on spatial economics, geographical location and spatial interactions between regions *matter*. Increasing returns to scale, together with enhancing market access, and probably a combination of labour migration across regions and vertical linkages between industries explain the cumulative process of regional growth, which endogenously turns into a polarisation of the spatial distribution of per capita income.[7] Additionally, the existence of localisation and urbanisation economies, or knowledge spillovers, reinforces the capacity of areas surrounding

[7]From the pioneering work of Marshall (1890) to the more recent developments of the "new economic geography" Krugman (1991, 1993), economists have emphasized a combination of these forces to explain the strong localization of economic activity.

more highly developed regions to grow. Not only geographical location but also proximity matters for growth. As can be seen in Table 1, the Chinese provinces with higher growth rates of per capita GDP between 1978 and 2005 (over 9%) are all next to each other and stretch almost continuously from Liaoning in the northeast to Guangdong in the southeast and to Sichuan in the west. Liaoning, Hebei and Shandong are provinces located around Beijing and Tianjin; Zhejiang is located between Shanghai and Fujian; while Guangdong and Fujian are located next to Hong Kong, Taiwan and Shanghai. This set of coastal provinces also stretches westward through Anhui, Henan, Hubei and Sichuan. Proximity, or even the neighbourhood, could become a key factor in its growth.

In order to elucidate the existence and magnitude of these spatial spillovers, we conducted an analysis which hinges on the comparison of two income series: (i) *state-relative* income, where we normalise each province's per capita income by the per capita income in China (which are the data used to conduct the analysis in the previous subsections); (ii) and *neighbour-relative* income, where we normalised each province's per capita income by the average per capita income of the surrounding, physically contiguous provinces, excluding the province itself. As indicated by Quah (1996c), it is convenient to consider these two relative income series as the parts unexplained by nation-state factors and physical-location factors, respectively. The same analyses as those presented in subsections 3.1–3.3 to this new series of neighbour-relative income, focusing on the comparison with the state-relative income series. Interpretations are also straightforward: the closer the values of the neighbour-relative series are to unity, the lower the inequalities among neighbour provinces will be and, therefore, the higher the magnitude of the spillover effects will also be. From this, it can easily be inferred that comparing these two series would be equivalent to analysing *unconditional* versus *conditional* convergence.

4. Results

Results concerning both transition probability matrices and ergodic distributions are reported in tables 2 through 5. They constitute a total of 12 panels in which different sorts of related information are reported. The four different tables present results for the different sub-periods considered, i.e. the first panel in each table provides results for the entire 1952–2005 period, whereas the second and third panels provide results for the periods 1952–1978 and 1978–2005 respectively. Furthermore, apart from the analysis of the unweighted distribution of per capita income, the additional conditioning schemes commented on in the introduction are also considered (GDP-weighted, population-weighted and physically-contiguous conditioned). In addition, the last three rows in each panel display information on the initial, final and ergodic distribution of the variable under analysis.

4.1. Unweighted analysis

Table 2 reports on unweighted transition probability matrices for all the periods considered. In each of the matrices contained in the table, the upper limits have been set to the same values in order to make comparisons possible. The criterion to specify the grid is the one usually found in the literature (see, for instance Lamo, 2000), i.e. considering all observations for the entire period 1952–2005 (28 observations per year and 54 years, which totals 1512 observations), we divide them into five equally-sized intervals. The limits of the grid are displayed in the first row of each panel. Their interpretations are straightforward: the upper limit for the first state is 0.919, indicating that approximately one fifth of the total number of observations lay below that threshold—i.e. below 91.9% of the average. At the other extreme, the upper-state has observations lying above 1.061 (106.1%) of the average. Note that the average is unity, since our data have been normalised by the mean: the closer a value is to the unity, the closer it is to the average for its particular year.

The contents of each matrix in Table 2 have some commonalities with the concept of β-convergence, since they provide information about intra-distribution mobility, or churning (Quah, 1996c). Each cell in the matrices must be interpreted as the probability of remaining in that particular state after five years (recall that we compute 5-year transitions). For instance, the upper-left entry of the matrix in Table 2.a would indicate that the probability of the observations in the lowest relative per capita income state (below 0.919) remaining in that state was 82%, whereas the remainder moved up to higher relative income states. Persistence was even higher at the other extreme of the distribution, as revealed by the lower right cell in the matrix, which indicates that 89% of the observations in the highest relative income state remain in the same class after five years. The other values in the main diagonal show a higher degree of mobility. For instance, entry a_{22} would indicate that, after 5 years, only 62% of observations remain in the same state of relative wealth, whereas 15% move down to lower per capita income states and the remaining 23% move up to higher per capita income states. In general, values in the main diagonal closer to 1 indicate more persistence, whereas values closer to zero indicate higher mobility.

In the matrices examined in Table 2, values on the main diagonal average 0.72, 0.62 and 0.81 for the periods 1952–2005, 1952–1978 and 1978-2005, respectively. This information is rich, but it would be richer still if additional ways of evaluating persistence/mobility such as the mobility indices presented in Equation (7) were considered. The results of these indices are shown in Table 6 and, in general, they corroborate what the averages for the diagonal entries revealed, i.e. the sub-period 1952–1978 shows much higher mobility than that of 1978–2005 (0.737 vs. 0.605) and, for the whole period, total mobility lies somewhere in-between (0.674). However, as will be shown below, it is not only the intensity of mobility that differs across periods but, more importantly, its sign; mobility leads to probability mass concentrating at lower states in the first sub-period, whereas the pattern is the opposite for the second one.

The last three rows in each table support this claim. They contain information on the initial (1952), final (2005) and ergodic (steady-state) distributions for the three periods considered. Table 2.a indicates that the initial and final distributions do not differ strongly. What is more revealing is that, under current trends, the ergodic distribution will be almost uniform, with the same probability in each state. However, we must bear in mind that in our particular setting under current trends may be a misleading statement, since trends have differed remarkably before and after the reform. The ergodic distributions in Table 2.b (pre-reform) and Table 2.c (post-reform) differ notably, not only compared to the ergodic distribution in Table 2.a but more notably with respect to each other. For the pre-reform period (1952-1978) the ergodic distribution skews towards the left tail of the distribution, whereas the opposite is found for the post-reform period (1978-2005). This would imply that the effects of the reform were positive for convergence among provinces and they are likely to continue over time, indicating that, under 1978-2005 trends, the two states of highest relative per capita income will contain 64% of the provinces.

It is relevant not only to compute the values of the steady-state distribution but also to analyse the speed at which it is approached. As indicated in previous sections, this can be evaluated via the concept of the asymptotic half life of the chain, i.e. the amount taken to cover half the distance from the ergodic distribution (Magrini, 1999). Therefore, computing Equation (6) leads to the results in Table 7. As one might expect *a priori*, although the steady-state reached considering only 1978–2005 information (Table 2.c) is more favourable than that obtained using 1952–1978 information (Table 2.b), it will take much longer to reach the former, in fact, virtually twice the time. This is the result of the higher intra-distribution mobility found for the pre-reform period, as revealed by Table 6. Therefore, although the future predicted using only 1978–2005 information is far more promising, it will take longer to reach it.

Bulli (2001), Johnson (2000, 2005) and many others have pointed out that it may be problematic to consider a discrete approach in which probability is split in some states whose limits are somehow arbitrary. Sakamoto and Islam (2008) partly circumvent this criticism and add some additional robustness to their analysis by considering different grids (5 and 7 grids), the results being similar for the different choices. We believe it is more interesting to consider a *fully* continuous counterpart to the initial and final distributions in Table 2, but continuous counterparts to the steady-state distributions reported are also taken into account.

Figure 1.a displays the continuous counterparts to the discrete initial (1952) and final (2005) distributions in the tables corresponding to the unweighted analysis. Although the densities basically corroborate the results of the discrete analysis, we can perceive more clearly that, although convergence has taken place (the 2005 density is higher), we can also see that by 1978 the distribution became bimodal. Therefore, there are some provinces whose performance in terms of per capita income was much better than the rest. Figure 5.a displays the continuous counterpart to the steady-state distributions in tables 2.b and 2.c. Although results are generally corroborated, some subtleties that the 5-grid analysis could not show are perceived. These are basically related to the multi-modality that will prevail regardless of the sub-period that is considered to construct the steady-state distribution. Taking into account the pre-reform information, the ergodic density (solid line in Figure 5.a) would be basically unimodal, but some very rich provinces (upper tail of the distribution) will coexist with some others (fewer) that are very poor (lower tail of the distribution). This extreme behaviour will fade away if only post-reform information (dashed line in Figure 5.a) is considered, although we can still distinguish two bigger modes - twin peaks, to use Quah's (1996d) term for them.

Therefore, the results obtained by Sakamoto and Islam (2008) are generally corroborated, but we have complemented them in several ways. Although their way of normalising differs, they use a slightly shorter time period (1952–2003) and they add some robustness to the analysis by considering a different number of grids, we find the same broad results, i.e. divergence before the reform and strong convergence afterwards. However, the mobility indices, transition path analysis and continuous approach to the steady-state distributions all enrich the analysis.

4.2. Weighted analysis

The analysis performed in the previous section is relevant, but it might be judged as being partly biased because the same importance is attached to all provinces, especially if we are concerned about human welfare—the different provinces have different population sizes. As indicated in the introduction, the unweighted analysis could be highly misleading if we drew national borders differently, as this would affect the shape of the densities. It may be more natural to attach a weight to the observations, where the weights reflect the contribution of each observation in the sample. As indicated in previous sections, we will consider different weighting schemes, i.e. population and economic size (GDP). In the case of countries, both variables are very unevenly distributed. This is especially blatant in the case of population, for which India and China, two of the poorest countries in terms of per capita income, account for more than one third of the total population in the world, whereas some of the richest countries, such as Iceland or Luxembourg, account for only 0.01% of the world population (Goerlich, 2003). In our particular case, it does not seem fair either to treat all Chinese provinces equally in the estimation. As can be seen on the right-hand side of Table 1, there is a significant dispersion in the population of the different provinces. More important still, there has been an important dispersion in the growth rates of provincial population throughout the period analysed, with coefficients of variation in the average growth rates of the provincial population between 0.525 in the pre-reform period and 0.966 in the post-reform era. These differences and changes in the distribution of provincial population have relevant implications when we are looking at per capita distribution of GDP from an individual or personal welfare perspective instead of from a provincial point of view. For example, by 2005, as indicated in Table 1, the population in Sichuan was 110,060,000 (larger than any European country), whereas that of Qinghai was 5,430,000. Therefore, the welfare implications of Sichuan converging with the rest of the provinces are not the same as if Qinghai converged, because of the number of people involved.

Results are shown in the GDP-weighted and population-weighted panels in Tables 3 and 4 respectively. The mobility indices, transition path analysis and continuous analysis are reported in the same tables and figures as those corresponding to the unweighted analysis. Both tables 3 and 4 offer new perspectives on the evaluation of convergence. Although the unweighted analysis did not predict convergence or divergence (in accordance with the ergodic distribution), the weighted analyses did yield different results. Under the population-weighted scenario (Table 4), according to which we evaluate transitions of people moving across classes, considering the entire period (1952–2005), the steady-state distribution has more than half of the probability mass (56%) in the two upper states. This indicates that a large part of the *population* will escape from poverty in the long run. However, similarly to what we obtained for the unweighted analysis, the tendencies differ remarkably between the pre- and post-reform sub-periods, and it is the effect of the second sub-period which drives the convergence pattern most. As shown in Table 4.b, although the predicted pattern using the 1952–1978 information was convergence, most of the population was being driven deep down into poverty, since the probability mass is overwhelmingly accumulated (72%) in the lowest relative per capita income states. This result is shared when weighting by economic size (Table 3.b), i.e. the largest provinces in terms of GDP were becoming relatively poorer. In contrast, the post-reform period shows opposite patterns. As Table 4.c reveals, in the hypothetical long run (i.e. under 1978–2005 trends) most of the population (94%) will reach the two highest per capita income states, and only 2% will remain in the poorest class. An analogous result is found when weighting by GDP (94% probability in the two wealthiest states), thus also indicating that large provinces in terms of GDP are also the ones that are escaping from poverty.

Although the general tendency when weighing by GDP or by population is similar for both sub periods, differences persist when evaluating the implied mobility in each matrix (Table 6) or the half-life time of convergence (Table 7). Regarding the latter, results are very similar to the unweighted case, while for the former some differences emerge. It is when weighting by GDP that convergence is faster in the post-reform sub-period, whereas this occurs in the first sub-period when weighting by population.

Finally, the continuous analysis in figures 1 and 5 further corroborates how relevant it is to perform the weighted analysis. Figure 2 and Figure 3 report information already displayed in Figure 1 in a different way so as to facilitate an easier visualisation of the patterns. As shown in Figure 1.b and Figure 1.c and, particularly, in figures 2 and 3, the evolution of the shape of the *weighted* densities differs compared to the one shown in Figure 1.a, especially in 2005. Weighting by GDP makes the density shift rightwards (Figure 2.c), although some additional bumps emerge, thus indicating that some important shares of GDP will remain in poor provinces. The result of weighting by population is more striking, since it indicates that by 2005 a large share of the population was reaching higher income levels, but a larger share was also trailing behind, as indicated by a marked bimodality. This is what Quah (1996d) refers to as "twin peaks". However, in the steady-state (figures 5.b and 5.c), and confirming what we found via the discrete analysis of the transition matrices, much of this bimodality will fade away, and the distributions will be basically skewed rightwards when using 1978–2005 information, which contrasts sharply with the bimodality found for the unweighted case (Figure 5.a).

In synthesis, uneven distribution of per capita GDP across Chinese provinces becomes less strong when weighted by GDP or population, that is, in terms of average personal welfare, and when using the post-reform information the implicit steady state distributions will be skewed rightwards and reflect an improvement in the symptoms of convergence. Nevertheless, some peaks persist on the upper tails of the distribution and it will also take a long time to reach the steady state. These stylised facts, together with the variability and changes of the trend in the variance of the distribution mentioned above, are quite consistent with the timing of the reforms, the unbalanced regional implications of these reforms and with the changes of emphasis in the main policy objectives during the period.

In the first phase of the economic reforms, but before economic liberalisation, the strategy was concentrated on the rural areas. The commune system was removed in favour of the Household Responsibility System, where workers were allowed to operate on their own, although with some restrictions.[8] After decollectivatisation, the Chinese government promoted economic policies addressed to diversify agriculture, especially by enhancing the rural industries and the township and village enterprises (TVEs). In fact, the promotion of TVEs was the most important way to transfer excess rural labour into industrial production, given the strong restrictions on interprovincial migration (Fujita and Hu, 2001). As a result, rural industrial output increased sharply in this period. However, the effectiveness of TVEs also raised some doubts owing to the fact that they often operated according to non-economic criteria in the early years of the reforms. Some regions improved in this phase, especially those oriented towards industry, but the income differentials persisted among provinces given the barriers that existed across provinces (Rozelle, 1994).

In the 1980s, the second phase of the reforms was characterised by the gradual opening up of the Chinese economy, the increased presence of the non-state sector (collective and private sectors) and a fiscal reform that endowed the provinces with more fiscal power (Wei, 1996). At first the open-door policy was especially favourable for the coastal areas (open cities and Special Economic Zones —SEZs—). Thus the geographic and economic policy factors allowed trade and FDI to become concentrated in the coastal areas.[9] At the same time, this period was distinguished by a major liberalisation and decentralization of the economy compared with the previous stage. For example, price liberalisation accelerated the entry of non-state enterprises, and the profit-oriented incentive schemes in state industry led to a rapid increase in industrial output and gains in productivity by the mid-1980s. As a result, the non-state sector gradually became more important in the economic development of China.[10] Although the interprovincial mobility of workers was still costly, there was an increase in migrational movements from rural areas to urban and coastal areas. On the other hand, the fiscal decentralisation of 1980 played a key role in improving the autonomy of local governments, but generated a significant budget deficit. Consequently, the fiscal system was reformed in 1985. The immediate effect of this reform was a reduction in the central government's ability to redistribute revenues among regions which, together with the economic developments that favored coastal provinces, increased symptoms of divergence and led to a new fiscal reform in 1994. The main feature of this reform was the separation of the national tax service from the local tax service, with an additional mixed category that was shared between central and local government, without negotiation and applied to all the provinces with the aim of reducing the income gap across provinces.

In 1995, the Chinese government recognised that:

"Since the adoption of reforms and open-door policies, we have encouraged some regions to

[8] Further details on rural reforms and agricultural growth can be found in Lin (1992).

[9] Although FDI was allowed in 1979, the effects on output are more significant in the 1980s and 1990s.

[10] Further details about the effects of the reform on the performance of the Chinese State Enterprises can be found in (Li, 1997).

develop faster and get richer, and we have advocated that the richer should act as a model for and help the poor. Each region has had immense economic development and the people's standard of living has had great improvement. But for some reasons, regional economic inequalities have widened somewhat" (People's Daily Overseas Edition, October 5, 1995, p. 4.)

Thus, the strategy was changing in favour of promoting a more evenly balanced regional development, in an attempt to reduce the tendencies towards uneven regional development. This strategy became evident in the Eighth Five-Year Plan and, more especially, in the Ninth Five-Year Plan (1996–2000). The Chinese government launched a strategy to promote the development of the central and western regions relied, at least partly, on the spillovers generated by the more developed coastal provinces.

4.3. Conditioning: spatial analysis

The transition probability matrices in Tables 5 show neighbour-relative counterparts to the transition probability analysis carried out for weighted and unweighted state-relative series (tables 2, 3 and 4). Likewise, the neighbor-relative analysis indicates that conclusions differ notably prior to and after the reform, i.e. they hinge critically on whether we base the future projections (ergodic distributions) on 1952–1978 or 1978–2005 information.

If the entire period 1952–2005 (Table 5.a) is considered, the diagonal entries average to 0.708, which is lower than the 0.718 corresponding to the state-relative series (Table 2.a). The mobility indices in Table 6 corroborate this finding, since $\mu_1 = 0.674$ in the case of the state-relative series, and $\mu_1 = 0.695$ when conditioning by neighbours' information. Under these trends, the (slightly) higher mobility would lead to an apparently multi-modal ergodic distribution, but it is difficult to discern tendencies. The analysis for the different sub-periods shows, once more, different patterns. The ergodic distribution corresponding to the 1952–1978 trends indicates that multi-modality will prevail in the future. Multi-modality vanishes if we focus on 1978–2005 trends (Table 5.c). In both cases, but especially for 1978–2005 trends, the ergodic distribution differs remarkably when compared to state-relative information (tables 2.c, 3.c and 4.c), since the probability mass does not entirely abandon the central states. However, as indicated by the asymptotic half life of convergence in Table 7, the ergodic distribution will be achieved much faster when conditioning by neighbouring information, i.e. *conditional* convergence will be faster.

Figure 1.d and Figure 4 also show the impact of conditioning on neighbouring-province information. Although the information contained in Figure 4 was already reported in Figure 1.a and Figure 1.d, the way it is presented allows a clearer understanding of the effect of spatial conditioning. Both Figure 1.d and Figure 4 show tighter distributions for neighbour-relative compared to state-relative per capita income series. This would indicate that each province's per capita income is closer to the average of its surrounding provinces than to the national average, thereby suggesting that spatial spillovers do matter. Yet some subtleties also exist. For instance, the uni-modal state-relative distribution of per capita income turns into a tighter but multi-modal distribution when conditioning by neighbouring information (Figure

4.c). This implies that, although the general tendency is towards convergence *within* spatial clusters, there are some provinces which outperform their neighbours, constituting a remarkable mode in the vicinity of 1.2 (Figure 4.c). Figure 4.c also shows how misleading it may be to draw conclusions based on summary statistics only. The implicit standard deviation of the state-relative series in Figure 4.c is 0.087, whereas that of the neighbour-relative series is higher (0.090), thus indicating more dispersion and, in principle, a flatter distribution. However, the driving force of the higher dispersion that is found is the increasing multi-modality. Therefore, spatial spillovers are relevant but *not* for everyone.

Figure 5.d reports continuous counterparts for the ergodic distributions in tables 5.b (solid line) and 5.c (dashed line). The solid line in Figure 5.d, corresponding to 1952–1978 trends, shows several modes, the biggest one in the vicinity of 1, but another two at the tails of the distribution. The dashed line indicates that the ergodic distribution that would prevail under 1978–2005 trends would be much tighter in the vicinity of 1, thus indicating that the members of spatial clusters' will be quite similar in terms of per capita income. However, several modes will lie in the upper tail of the distribution, thus indicating that, in the hypothetical long-run scenario, some provinces will still outperform their neighbours, i.e. although there will still be inequalities that cannot be explained by physical-location factors, they will affect provinces differently.

These results, and especially the tendency towards the stratification of provinces in different clubs, are of no minor concern to authorities, and reveal that there is still some room for policies promoting convergence in per capita GDP among Chinese provinces, because the natural tendency to spatial agglomeration seems to be persistent. Thus, together with the explicit regional policies and the use of other central government policies to re-balance regional development (central investment projects, endowment of infrastructures, credit policy, etc.), other measures are also needed to balance the tendency towards the localisation of economic activity induced by market forces. Improvements in the accessibility and the role of market mechanisms in the interior are needed, but increasing the role assigned to official interprovincial migrations is probably necessary too.

5. Conclusions

Nobody doubts that the acceleration of the economic reforms initiated at the end of the 1970s have encouraged economic growth over the last four decades. The open-door policy, with a strong drive towards industrialisation focused on foreign investment, especially in the coastal regions, along with a series of economic reforms oriented more towards the market probably explained this exceptional performance during the 1980s and 1990s. In 1995, however, the Chinese government recognised that the income gap between western and central regions and the coastal areas was increasing, thus making it necessary to implement pro-active policies to reduce these inequalities. The stimulus package that was carried out was focused mainly on the development of inland provinces through the promotion of investment as a way to reduce those imbalances.

Accordingly, a plethora of research studies have examined not only the aggregate growth of the country but also other related questions, such as whether differences in per capita income across provinces exist, along with the evolution of disparities over time. This ample body of literature analysing convergence across Chinese provinces continuous to grow, examining such relevant topics as those examined by the country and regional convergence studies. Some papers have analysed provincial convergence following the early proposals of Barro and Sala-i-Martin (1992), i.e. by examining β- and σ-convergence, together with some of the ulterior refinements of these techniques. Some others (fewer) have leaned towards the distribution dynamics' model initially proposed by Quah (1993a,b). Our article follows this second line of research. Recent contributions such as Bhalla et al. (2003) or Sakamoto and Islam (2008), have applied Quah's basic proposals to examine provincial convergence in per capita income. Our paper complements their methods and findings and extend them in several directions.

Similarly to Sakamoto and Islam (2008), the ergodic distributions obtained using either pre-reform or post-reform information are quite different, a positively skewed distribution being produced for 1952–1978 and a negatively skewed one for the period 1978–2005. Therefore, it would be corroborated that the post-reform policies have led most provinces to escape from relatively low per capita income levels. However, this analysis has some limitations such as the need to specify a *discrete* grid with a limited number of states. Few contributions try to fix this by considering a *continuous* state space approach (Johnson, 2000). We follow Johnson's (2000) approach to provide continuous counterparts to the ergodic distributions yielded by transition probability matrices, which offered more detailed results. Under both pre- and post-reform information, the hypothetical long-run scenario shows multi-modality. For the 1952–1978 information, the distances separating the biggest modes are quite large, with predominance of a large mode comprising most of the provinces with incomes close to the average, and a small group of provinces that are becoming very rich. However, using 1978–2005 information these two modes become more balanced, with one of them above the national average and the other one below it. We can also corroborate Sakamoto and Islam's claim that "the dynamics of the post-reform period do not yet seem to have settled into a stable pattern". In our case, the analysis of the asymptotic half life of convergence indicates that it will take much longer to reach the steady state under 1978–2005 trends. Under this scenario, although most provinces will escape from relative poverty, it will take longer because of more complex intra-distribution dynamics.

We extend the analysis to control for some relevant characteristics of Chinese provinces. Specifically, although *unweighted* analysis of country/regional convergence is commonplace, *weighted* analysis is far less widely extended. However, in many circumstances and especially if we focus on human welfare (Sala-i-Martin, 2006), weighted analysis might be more relevant than its unweighted counterpart. Several weighting schemes are possible, but because of their significance we considered the population and economic size (GDP) of each province. As forcefully stressed throughout the article, since both population and GDP differences across Chinese provinces are outstanding, controlling for these differences might alter the results substantially—which in fact turned out to be the case.

For the entire period 1952–2005, we find that under the population-weighted scenario, the steady-state distribution has more than half of the probability mass (56%) in the two upper states, thereby indicating that much of the population will escape from poverty in the long run. As expected, the tendencies differ remarkably between the pre- and the post-reform periods, and it is the effect of the second sub-period (1978-2005) which, for the most part, drives the convergence pattern. Specifically, for the pre-reform period, although the predicted pattern was convergence, most of the population was driven deep down into relative poverty, since probability mass is overwhelmingly accumulated (72%) in the lowest relative per capita income states. This result is shared when weighting by economic size. However, the continuous ergodic distributions also indicate that some provinces will still be much richer than the rest, as indicated by the existence of several bumps well above the unity. In contrast, the post-reform period, shows opposite patterns. In the hypothetical long run (under 1978-2005 trends) most of the population (94%) will reach the two highest per capita income states, and only 2% would remain in the poorest class. An analogous result is found when weighting by GDP (94% probability in the two wealthiest states), thus indicating that large provinces in terms of GDP are also the ones escaping from poverty. Moreover, when weighting by GDP, convergence is faster in the post-reform period, whereas this occurs in the first sub-period (1952-1978), when weighting by population. Thus, the marked bimodality yielded by the unweighted analysis turns into a tighter pattern of convergence when weighting by population of each province or economic size, as suggested by Sala-i-Martin (2006).

Finally, our study also analysed whether spatial spillovers exist. Although a more thorough analysis would be welcome, the techniques we use can be easily adapted to provide some insights into the magnitude of these effects. This can be thought of as conditional convergence analysis, in which a province is compared only with its contiguous provinces, and therefore it could converge towards its neighbours' average (conditional, club or cluster convergence) instead of towards the national average (unconditional convergence). Compared to the unweighted, unconditional analysis, the long-run scenario will be multi-modal under pre-reform trends—in fact, we could even talk of club *divergence*. However, under post-reform trends, cluster convergence will be much stronger, with probability mass concentrating tightly around unity as indicated by the ergodic density; provinces will converge strongly with their neighbours, although some amount of multi-modality will still prevail.

According to our results it seems that all the reforms have enabled poorer regions to gradually converge with the richer ones in the post-reform period, while in the pre-reform period no convergence was found. However, more economic reforms are needed in this regard to guarantee balance and steady economic growth, thereby improving the standards of living of the whole population.

References

Alesina, A. and Perotti, R. (1993). Income distribution, political instability, and investment. Working Paper 4486, NBER.

Alesina, A. F. and Rodrik, D. (1994). Distributive politics and economic growth. The Quarterly Journal of Economics, 109:465–490.

Bai, C. E., Hsieh, C., and Qian, Y. (2006). The return to capital in China. Technical Report 12755, NBER.

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

Bhalla, A., Yao, S., and Zhang, Z. (2003). Regional economic performance in China. The Economics of Transition, 11(1):25–39.

Bramall, C. (2000). Sources of Chinese Economic Growth, 1978–1996. Studies on Contemporary China. Oxford University Press.

Brun, J. F., Combes, J. L., and Renard, M. F. (2002). Are there spillover effects between coastal and noncoastal regions in China? China Economic Review, 13(2-3):161–169.

Bulli, S. (2001). Distribution dynamics and cross-country convergence: a new approach. Scottish Journal of Political Economy, 48:226–243.

Caselli, F., Esquivel, G., and Lefort, F. (1996). Re-opening the convergence debate: a new look at cross-country growth empirics. Journal of Economic Growth, 1(3):363–389.

Cetorelli, N. (2002). Could Prometheus be bound again? A contribution to the convergence controversy. Journal of Economic Dynamics and Control, 27(1):29–50.

Chai, J. C. (1998). China: Transition to a Market Economy. Studies on Contemporary China. Oxford University Press.

Chen, J. and Fleisher, B. (1996). Regional income inequality and economic growth in China. Journal of Comparative Economics, 22(2):141–164.

Chow, G. (2006). Are Chinese official statistics reliable? CESifo Economic Studies, 52(2):396–414.

De la Fuente, Á. (1997). The empirics of growth and convergence: A selective review. Journal of Economic Dynamics and Control, 21(1):23–73.

Ding, L., Haynes, K., and Liu, Y. (2008). Telecommunications infrastructure and regional income convergence in China: panel data approaches. The Annals of Regional Science, 42(4):843–861.

Durlauf, S. N. and Quah, D. T. (1999). The new empirics of economic growth. In Handbook of Macroeconomics, volume 1A, pages 231–304. North Holland, Amsterdam.

Fujita, M. and Hu, D. (2001). Regional disparity in China 1985–1994: The effects of globalization and economic liberalization. The Annals of Regional Science, 35(1):3–37.

Geweke, J., Marshall, R. C., and Zarkin, G. A. (1986). Mobility indices in continuous time Markov chains. Econometrica, 54(6):1407–1423.

Goerlich, F. J. (2003). Weighed, kernel density estimators and convergence. Empirical Economics, 28(2):335–351.

Gundlach, E. (1997). Regional convergence of output per worker in China: a neoclassical interpretation. Asian Economic Journal, 11(4):423–442.

Härdle, W. and Linton, O. (1994). Applied nonparametric methods. In Engle, R. and McFadden, D., editors, Handbook of Econometrics, volume 4. North Holland, Amsterdam.

Holz, C. A. (2005). China governance project: The institutional arrangements for the production statistics. Statistics Working Papers 2005/1, OECD, Paris.

Islam, N. (2003). What have we learnt from the convergence debate? Journal of Economic Surveys, 17(3):309–362.

Jian, T., Sachs, J., and Warner, A. (1996). Trends in regional inequality in China. China Economic Review, 7(1):1–21.

Johnson, P. A. (2000). A nonparametric analysis of income convergence across the US states. Economics Letters, 69:219–223.

Johnson, P. A. (2005). A continuous state space approach to "Convergence by parts". Economics Letters, 86:317–321.

Jones, C. I. (1997). On the evolution of the world income distribution. Journal of Economic Perspectives, 11(3):19–36.

Kremer, M., Onatski, A., and Stock, J. (2001). Searching for prosperity. Carnegie-Rochester Conference Series on Public Policy, 55:275–303.

Krugman, P. R. (1991). Increasing returns and economic geography. Journal of political economy, 99(3):483.

Krugman, P. R. (1993). Geography and Trade. MIT Press.

Lamo, A. (2000). On convergence empirics: Some evidence for Spanish regions. Investigaciones Económicas, 24:681–707.

Lardy, N. R. (1992). Foreign Trade and Economic Reform in China, 1978–1990. Cambridge University Press.

Li, W. (1997). The impact of economic reform on the performance of Chinese state enterprises, 1980–1989. Journal of Political Economy, 105(5):1080–1106.

Lin, J. Y. (1992). Rural reforms and agricultural growth in China. The American Economic Review, pages 34–51.

Maasoumi, E. and Wang, L. (2008). Economic reform, growth and convergence in China. Econometrics Journal, 11(1):128–154.

Magrini, S. (1999). The evolution of income disparities among the regions of the European Union. Regional Science and Urban Economics, 29(2):257–281.

Marshall, A. (1890). Principles of Economics. Macmillan and Co.

Pedroni, P. and Yao, J. Y. (2006). Regional income divergence in China. Journal of Asian Economics, 17(2):294–315.

Quah, D. T. (1993a). Empirical cross-section dynamics in economic growth. European Economic Review, 37:426–434.

Quah, D. T. (1993b). Galton's fallacy and tests of the convergence hypothesis. Scandinavian Journal of Economics, 95(4):427–443.

Quah, D. T. (1996a). Aggregate and regional disaggregate fluctuations. Empirical Economics, 21:137–159.

Quah, D. T. (1996b). Empirics for economic growth and convergence. European Economic Review, 40:1353–1375.

Quah, D. T. (1996c). Regional convergence clusters across Europe. European Economic Review, 40:951–958.

Quah, D. T. (1996d). Twin Peaks: Growth and convergence in models of distribution dynamics. Economic Journal, 106(437):1045–1055.

Quah, D. T. (1997a). Empirics for growth and distribution: Stratification, polarization and convergence clubs. Journal of Economic Growth, 2(1):27–59.

Quah, D. T. (1997b). Regional cohesion from local isolated actions. II. Conditioning. Discussion Paper 379, CEPR.

Raiser, M. (1998). Subsidising inequality: Economic reforms, fiscal transfers and convergence across Chinese provinces. Journal of Development Studies, 34(3):1–26.

Redding, S. (2002). Specialization dynamics. Journal of International Economics, 58(2):299–334.

Rozelle, S. (1994). Rural industrialization and increasing inequality: Emerging patterns in China's reforming economy. Journal of Comparative Economics, 19(3):362 – 391.

Sakamoto, H. and Islam, N. (2008). Convergence across Chinese provinces: An analysis using Markov transition matrix. China Economic Review, 19(1):66–79.

Sala-i-Martin, X. (2006). The world distribution of income: Falling poverty and... convergence, period. Quarterly Journal of Economics, 121(2):351–398.

Schultz, T. P. (1998). Inequality in the distribution of personal income in the world: How it is changing and why. Journal of Population Economics, 11(3):307–344.

Shorrocks, A. F. (1978). The measurement of mobility. Econometrica, 46(5):1013–1024.

Temple, J. (1999). The new growth evidence. Journal of Economic Literature, 37:112–156.

Wang, Z. (2003). Regional Divergence of per capita GDP in China: 1991-1999. Journal of Chinese Economic and Business Studies, 2:39–53.

Weeks, M. and Yao, J. Y. (2003). Provincial conditional income convergence in China, 1953–1997: A panel data approach. Econometric Reviews, 22(1):59–77.

Wei, Y. (1996). Fiscal systems and uneven regional development in China, 1978–1991. Geoforum, 27(3):329–344.

Yao, S. and Zhang, Z. (2001a). On regional inequality and diverging clubs: a case study of contemporary China. Journal of Comparative Economics, 29(3):466–484.

Yao, S. and Zhang, Z. (2001b). Regional growth in China under economic reforms. Journal of Development Studies, 38(2):167–186.

Table 1: Descriptive statistics for Chinese provinces, per capita income (GDP/N) and population (N), 1952 to 2005

Province	Y/N[a]			Y/N annual growth rate (%)			N[b]			N annual growth rates (%)		
East	**1952**	**1978**	**2005**	**1952-78**	**1978-05**	**1952-05**	**1952**	**1978**	**2005**	**1952–78**	**1978–05**	**1952–05**
Shanghai	430	2,944	23,583	7.68	8.01	7.85	573	1,098	1,778	2.53	1.80	2.16
Beijing	170	1,485	11,351	8.69	7.82	8.25	490	872	1,538	2.24	2.12	2.18
Tianjin	299	1,112	9,072	5.18	8.08	6.65	439	724	1,043	1.94	1.36	1.65
Liaoning	218	815	6,956	5.20	8.27	6.75	1,932	3,394	4,220	2.19	0.81	1.49
Jiangsu	131	310	5,792	3.37	11.45	7.41	3,739	5,834	7,468	1.73	0.92	1.31
Zhejiang	112	274	5,921	3.50	12.06	7.77	2,213	3,751	4,894	2.05	0.99	1.51
Guangdong	101	532	7,100	6.60	10.07	8.35	3,170	5,593	10,022	2.21	2.18	2.20
Shangdong	91	286	4,652	4.50	10.88	7.71	4,827	7,160	9,248	1.53	0.95	1.23
Fujian	102	234	4,195	3.25	11.28	7.26	1,270	2,446	3,535	2.55	1.37	1.95
Guangxi	67	202	1,662	4.34	8.12	6.25	1,943	3,402	4,660	2.18	1.17	1.66
Hebei	125	341	3,359	3.94	8.84	6.41	3,272	5,057	6,851	1.69	1.13	1.40
Mean	167.82	775.91	7,603.91	5.11	9.54	7.33	2,169.82	3,575.55	5,023.36	2.08	1.35	1.70
Median	125.00	341.00	5,921.00	4.50	8.84	7.41	1,943.00	3,402.00	4,660.00	2.18	1.17	1.65
Standard deviation	109.38	830.58	5,932.76	1.82	1.63	0.73	1,450.00	2,171.69	3,060.77	0.33	0.48	0.36
Central												
Heilongjiang	238	399	2,608	2.01	7.20	4.62	1,111	3,130	3,820	4.06	0.74	2.36
Jilin	153	324	3,237	2.93	8.90	5.93	1,065	2,149	2,715	2.74	0.87	1.78
Hubei	90	215	2,524	3.41	9.55	6.49	2,751	4,575	5,710	1.98	0.82	1.39
Shanxi	116	322	2,457	4.00	7.82	5.93	1,395	2,424	3,355	2.15	1.21	1.67
Hunan	86	211	1,809	3.51	8.28	5.92	3,271	5,166	6,326	1.77	0.75	1.25
Anhui	78	120	1,349	1.67	9.38	5.53	2,966	4,713	6,120	1.80	0.97	1.38
Jiangxi	114	179	1,658	1.75	8.59	5.18	1,656	3,183	4,307	2.54	1.13	1.82
Henan	83	157	1,905	2.48	9.69	6.09	4,371	7,067	9,380	1.87	1.05	1.45
Inner Mongolia	173	300	3,482	2.14	9.50	5.83	716	1,823	2,386	3.66	1.00	2.30
Mean	125.67	247.44	2,336.56	2.66	8.77	5.72	2,144.67	3,803.33	4,902.11	2.51	0.95	1.71
Median	114.00	215.00	2,457.00	2.48	8.90	5.92	1,656.00	3,183.00	4,307.00	2.15	0.97	1.67
Standard deviation	53.35	92.63	719.93	0.84	0.87	0.55	1,242.15	1,707.73	2,214.94	0.84	0.17	0.40
West												
Sichuan	171	488	5,402	4.12	9.31	6.73	6,405	9,707	11,006	1.61	0.47	1.03
Xinjiang	166	260	2,519	1.74	8.77	5.27	465	233	2,008	-2.62	8.30	2.80
Qinghai	101	376	2,003	5.19	6.39	5.80	161	365	543	3.20	1.48	2.32
Ningxia	126	403	3,006	4.57	7.73	6.17	142	356	595	3.60	1.92	2.74
Gansu	125	322	2,315	3.71	7.58	5.66	1,065	1,870	2,919	2.19	1.66	1.92
Shaanxi	85	322	2,314	5.26	7.58	6.43	1,528	2,779	3,718	2.33	1.08	1.69
Yunnan	70	178	1,379	3.65	7.88	5.78	1,695	3,091	4,442	2.34	1.35	1.83
Guizhou	123	248	1,563	2.73	7.06	4.91	1,490	2,686	3,730	2.29	1.22	1.75
Mean	120.88	324.63	2,562.60	3.87	7.79	5.84	1,618.88	2,635.88	3,620.13	1.87	2.19	2.01
Median	124.00	322.00	2,314.50	3.91	7.65	5.79	1,277.50	2,278.00	3,318.50	2.31	1.42	1.88
Standard deviation	35.59	97.86	1,259.11	1.20	0.92	0.60	2,031.27	3,092.53	3,315.74	1.92	2.51	0.59
Total 28 provinces												
Mean	140.86	477.11	4,470.46	3.97	8.79	6.39	2,004.32	3,380.29	4,583.46	2.15	1.46	1.79
Median	119.50	316.00	2,807.00	3.68	8.44	6.21	1,592.00	3,110.50	4,020.00	2.19	1.13	1.72
Standard deviation	78.07	566.90	4,494.13	1.71	1.40	0.99	1,538.61	2,308.62	2,856.81	1.13	1.41	0.45

[a] In 1952 yuan/person.
[b] In 10,000 persons.

Table 2: Transition probability matrix and ergodic distribution, per capita income (GDP/N), unweighted, 5-year transitions, limits all years

	Upper limit, all years:				
(Number of observations)	0.919	0.960	0.990	1.061	Max.
(282)	0.82	0.15	0.01	0.01	0.01
(274)	0.15	0.62	0.19	0.02	0.02
(287)	0.02	0.21	0.58	0.18	0.00
(280)	0.00	0.02	0.21	0.68	0.09
(276)	0.00	0.01	0.00	0.10	0.89
Initial distribution	0.18	0.21	0.11	0.29	0.21
Final distribution	0.14	0.25	0.14	0.21	0.25
Ergodic distribution	0.19	0.20	0.20	0.22	0.20

a) 1952–2005

	Upper limit, all years:				
(Number of observations)	0.919	0.960	0.990	1.061	Max.
(121)	0.68	0.27	0.03	0.02	0.00
(165)	0.18	0.57	0.22	0.02	0.01
(94)	0.09	0.34	0.33	0.25	0.00
(133)	0.00	0.04	0.17	0.68	0.11
(130)	0.00	0.01	0.00	0.14	0.86
Initial distribution	0.18	0.21	0.11	0.29	0.21
Final distribution	0.25	0.14	0.25	0.18	0.18
Ergodic distribution	0.26	0.28	0.14	0.19	0.14

b) 1952–1978

	Upper limit, all years:				
(Number of observations)	0.919	0.960	0.990	1.061	Max.
(142)	0.93	0.07	0.00	0.00	0.00
(90)	0.07	0.78	0.12	0.01	0.02
(180)	0.00	0.16	0.68	0.15	0.01
(128)	0.00	0.00	0.20	0.73	0.08
(132)	0.00	0.01	0.00	0.06	0.93
Initial distribution	0.25	0.14	0.25	0.18	0.18
Final distribution	0.14	0.25	0.14	0.21	0.25
Ergodic distribution	0.11	0.14	0.12	0.23	0.41

c) 1978–2005

Table 3: Transition probability matrix and ergodic distribution, per capita income (GDP/N), GDP-weighted, 5-year transitions, limits all years

	Upper limit, all years:				
(Share of GDP)	0.919	0.960	0.990	1.061	Max.
(0.14)	0.80	0.18	0.01	0.00	0.00
(0.21)	0.13	0.68	0.15	0.01	0.02
(0.17)	0.03	0.21	0.57	0.19	0.00
(0.21)	0.00	0.03	0.18	0.71	0.08
(0.27)	0.00	0.01	0.00	0.09	0.90
Initial distribution	0.13	0.19	0.08	0.36	0.23
Final distribution	0.06	0.13	0.06	0.26	0.49
Ergodic distribution	0.19	0.23	0.14	0.20	0.25

a) 1952–2005

	Upper limit, all years:				
(Share of GDP)	0.919	0.960	0.990	1.061	Max.
(0.15)	0.70	0.27	0.03	0.01	0.00
(0.24)	0.19	0.58	0.22	0.00	0.01
(0.17)	0.12	0.28	0.44	0.16	0.00
(0.19)	0.00	0.06	0.17	0.68	0.09
(0.25)	0.00	0.02	0.00	0.12	0.87
Initial distribution	0.13	0.19	0.08	0.36	0.23
Final distribution	0.16	0.19	0.21	0.17	0.27
Ergodic distribution	0.38	0.34	0.16	0.08	0.04

b) 1952–1978

	Upper limit, all years:				
(Share of GDP)	0.919	0.960	0.990	1.061	Max.
(0.12)	0.92	0.08	0.00	0.00	0.00
(0.19)	0.03	0.83	0.09	0.01	0.04
(0.17)	0.00	0.15	0.67	0.18	0.00
(0.23)	0.00	0.00	0.13	0.80	0.08
(0.29)	0.00	0.00	0.00	0.08	0.92
Initial distribution	0.16	0.19	0.21	0.17	0.27
Final distribution	0.06	0.13	0.06	0.26	0.49
Ergodic distribution	0.01	0.03	0.02	0.37	0.57

c) 1978–2005

Table 4: Transition probability matrix and ergodic distribution, per capita income (GDP/N), population-weighted, 5-year transitions, limits all years

(Share of population)	Upper limit, all years: 0.919	0.960	0.990	1.061	Max.
(0.28)	0.81	0.16	0.01	0.01	0.01
(0.24)	0.21	0.57	0.18	0.01	0.03
(0.19)	0.03	0.20	0.57	0.21	0.00
(0.18)	0.00	0.02	0.18	0.73	0.06
(0.13)	0.00	0.01	0.00	0.10	0.89
Initial distribution	0.22	0.28	0.09	0.21	0.20
Final distribution	0.15	0.21	0.09	0.28	0.26
Ergodic distribution	0.17	0.14	0.12	0.30	0.26

a) 1952–2005

(Share of population)	Upper limit, all years: 0.919	0.960	0.990	1.061	Max.
(0.26)	0.69	0.26	0.03	0.02	0.00
(0.37)	0.23	0.55	0.21	0.00	0.01
(0.16)	0.13	0.30	0.37	0.20	0.00
(0.11)	0.01	0.06	0.19	0.69	0.05
(0.10)	0.00	0.04	0.00	0.11	0.85
Initial distribution	0.22	0.28	0.09	0.21	0.20
Final distribution	0.33	0.10	0.25	0.20	0.12
Ergodic distribution	0.42	0.30	0.14	0.10	0.04

b) 1952–1978

(Share of population)	Upper limit, all years: 0.919	0.960	0.990	1.061	Max.
(0.28)	0.92	0.08	0.00	0.00	0.00
(0.11)	0.11	0.76	0.09	0.01	0.02
(0.21)	0.00	0.12	0.69	0.19	0.00
(0.24)	0.00	0.00	0.13	0.80	0.06
(0.16)	0.00	0.00	0.00	0.09	0.91
Initial distribution	0.33	0.10	0.25	0.20	0.12
Final distribution	0.15	0.21	0.09	0.28	0.26
Ergodic distribution	0.02	0.02	0.02	0.43	0.51

c) 1978–2005

Table 5: Transition probability matrix and ergodic distribution, per capita income (GDP/N), physically contiguous-conditioned, 5-year transitions, limits all years

	Upper limit, all years:				
(Number of observations)	0.919	0.960	0.990	1.061	Max.
(268)	0.79	0.17	0.03	0.01	0.00
(308)	0.15	0.64	0.17	0.03	0.01
(198)	0.05	0.24	0.48	0.21	0.02
(417)	0.00	0.04	0.12	0.76	0.07
(206)	0.00	0.01	0.01	0.11	0.87
Initial distribution	0.14	0.18	0.29	0.21	0.18
Final distribution	0.18	0.14	0.32	0.21	0.14
Ergodic distribution	0.22	0.23	0.14	0.26	0.15

a) 1952–2005

	Upper limit, all years:				
(Number of observations)	0.919	0.960	0.990	1.061	Max.
(107)	0.79	0.11	0.07	0.02	0.01
(119)	0.22	0.48	0.23	0.08	0.00
(91)	0.06	0.35	0.27	0.31	0.01
(227)	0.01	0.03	0.10	0.76	0.10
(99)	0.00	0.02	0.00	0.15	0.82
Initial distribution	0.14	0.18	0.29	0.21	0.18
Final distribution	0.21	0.29	0.07	0.29	0.14
Ergodic distribution	0.26	0.15	0.11	0.34	0.14

b) 1952–1978

	Upper limit, all years:				
(Number of observations)	0.919	0.960	0.990	1.061	Max.
(142)	0.79	0.19	0.01	0.01	0.00
(177)	0.09	0.80	0.11	0.00	0.00
(96)	0.00	0.10	0.73	0.14	0.03
(159)	0.00	0.02	0.13	0.86	0.00
(96)	0.00	0.00	0.00	0.01	0.99
Initial distribution	0.21	0.29	0.07	0.29	0.14
Final distribution	0.18	0.14	0.32	0.21	0.14
Ergodic distribution	0.04	0.20	0.24	0.26	0.26

c) 1978–2005

Table 6: Mobility indices (μ_1)[a]

Transition matrix	1952-1978	1978-2005	1952-2005
Unweighted	0.737	0.605	0.674
GDP-weighted	0.732	0.611	0.671
Population-weighted	0.755	0.614	0.670
Physically contiguous-conditioned	0.789	0.644	0.695

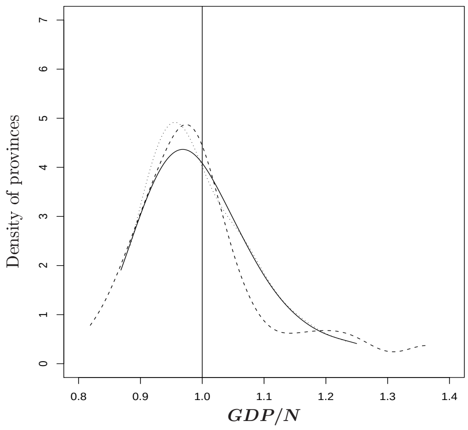
[a] See main text for definition of μ_1.

Table 7: Transition path analysis (asymptotic half life of convergence)[a]

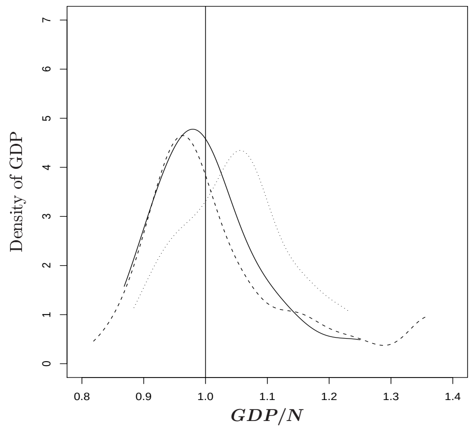
Transition matrix	1952–1978	1978–2005	1952–2005
Unweighted	16.795	29.820	15.044
GDP-weighted	52.627	22.414	19.541
Population-weighted	23.821	31.544	13.658
Physically contiguous-conditioned	9.762	15.218	11.609

[a] See main text for definition of $H - L$.

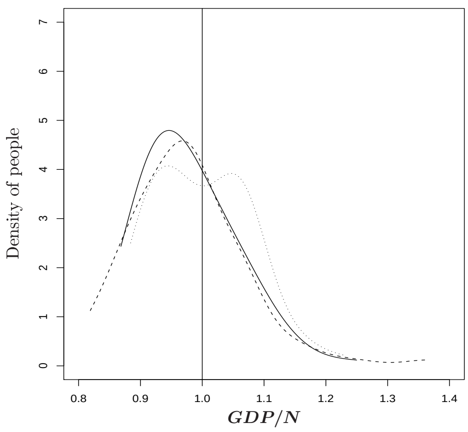
Figure 1: GDP/N, densities, 1952 vs. 1978 vs. 2005



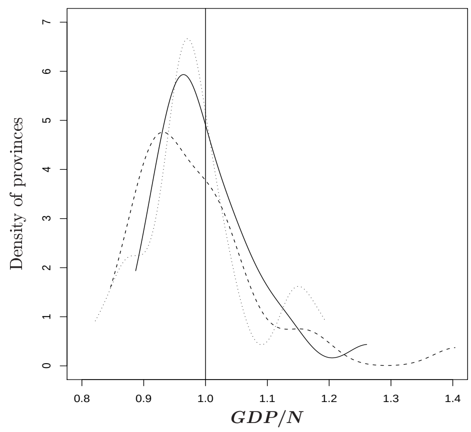
a) Unweighted



b) GDP-weighted



c) Population-weighted



d) Physically contiguous-conditioned

1952 —— 1978 ------ 2005 ········

Figure 2: GDP/N, densities, unweighted vs. GDP-weighted

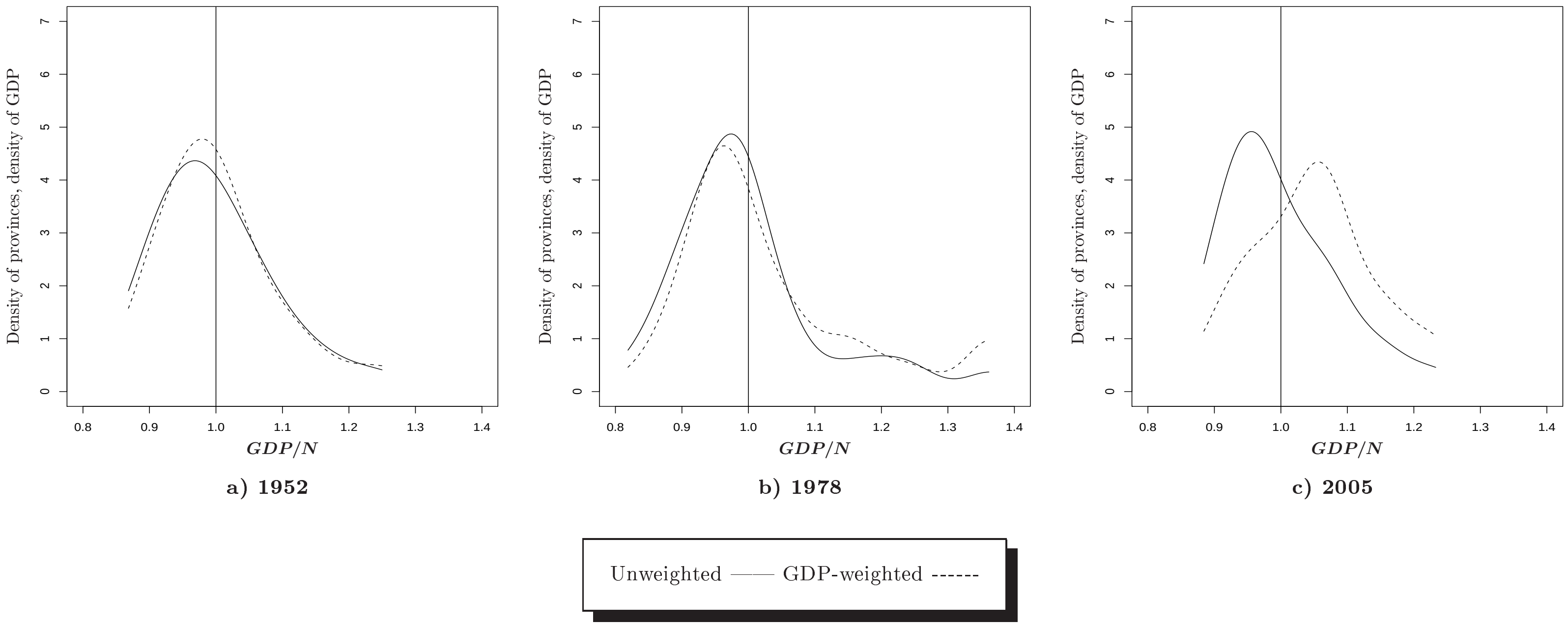
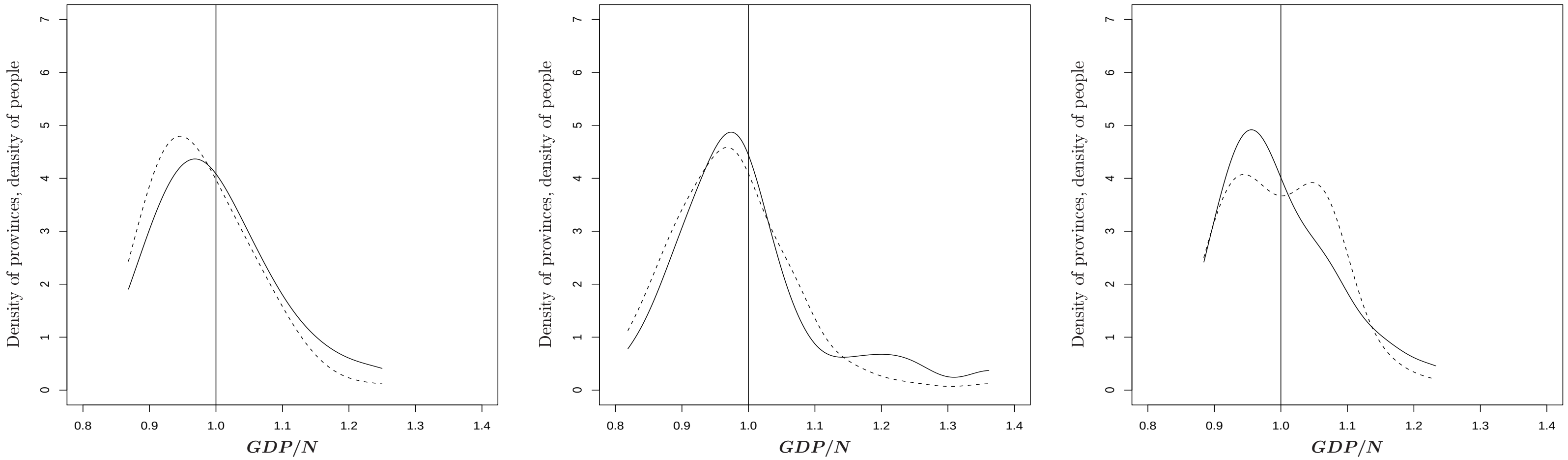


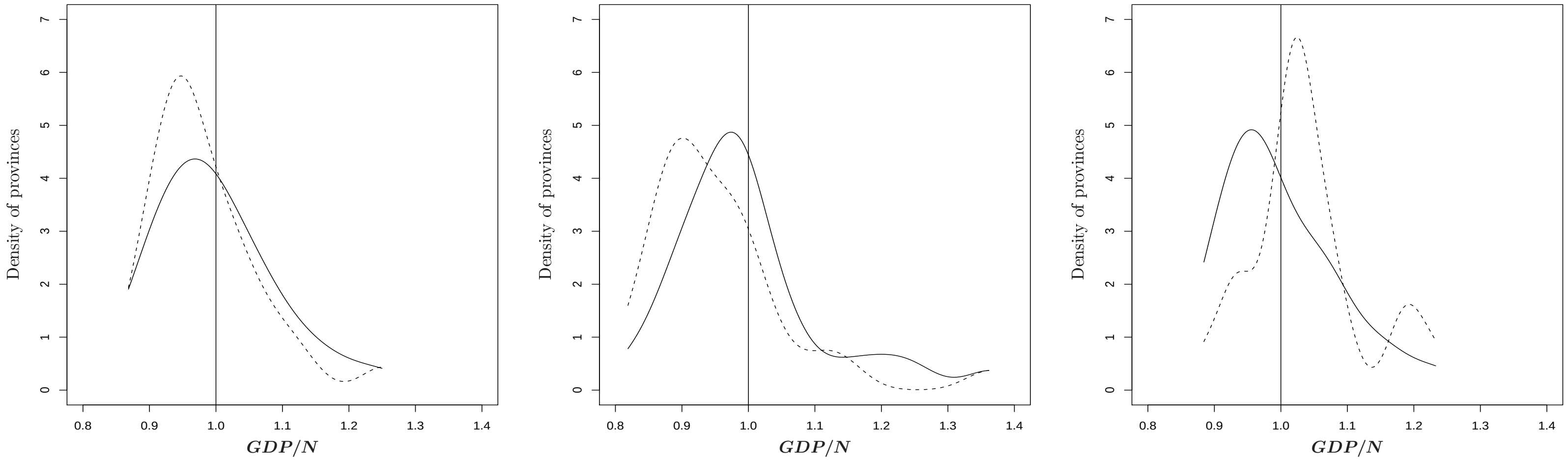
Figure 3: GDP/N, densities, unweighted vs. population-weighted



a) 1952 **b) 1978** **c) 2005**

Unweighted —— Population-weighted ------

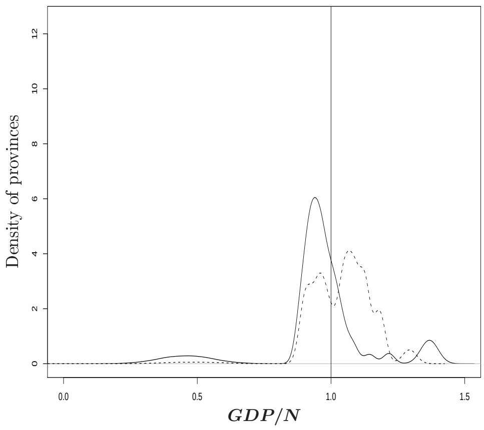
Figure 4: GDP/N, densities, unweighted vs. physically contiguous-conditioned



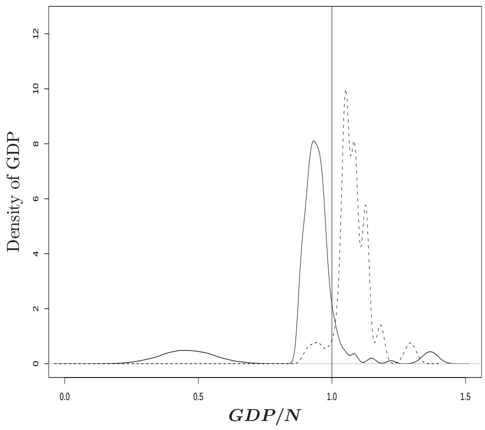
a) 1952 **b) 1978** **c) 2005**

Unweighted —— Physically contiguous-conditioned ------

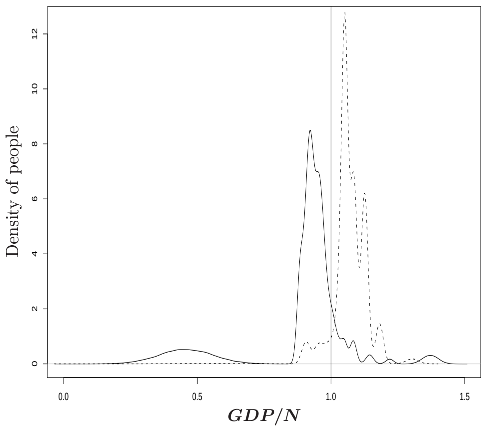
Figure 5: GDP/N, ergodic distributions, 1952–1978 vs. 1978–2005



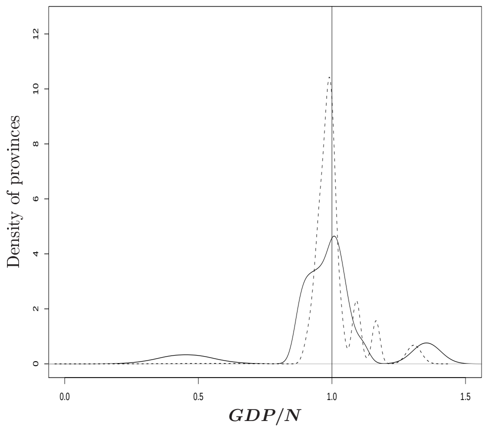
a) Unweighted



b) GDP-weighted



c) Population-weighted



d) Physically contiguous-conditioned

1952–1978 —— 1978–2005 ------