Income Convergence in the United States: A Tale of Migration and Urbanization

Authors	Riccardo DiCecio, and Charles S. Gascon
Working Paper Number	2008-002C
Revision Date	December 2008
Citable Link	https://doi.org/10.20955/wp.2008.002
Suggested Citation	DiCecio, R., Gascon, C.S., 2008; Income Convergence in the United States: A Tale of Migration and Urbanization, Federal Reserve Bank of St. Louis Working Paper 2008-002. URL https://doi.org/10.20955/wp.2008.002

Published In	Annals of Regional Science
Publisher Link	https://doi.org/10.1007/s00168-008-0284-1

Income Convergence in the United States: A Tale of Migration and Urbanization

Riccardo DiCecio*
Federal Reserve Bank of St. Louis

Charles S. Gascon
Federal Reserve Bank of St. Louis

December 3, 2008

* We are grateful to Costas Azariadis, Levon Barseghyan, Silvio Contessi, Cletus Coughlin, Mike Dueker, Kristie Engemann, George Fortier, Börje Johansson, Francesca Molinari, Howard Wall, and two anonymous referees for helpful comments. Any views expressed are our own and do not necessarily re.ect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System. All errors are our own.

Corresponding author: Riccardo DiCecio, FRB of St. Louis - Research Division, P.O. Box 442, St. Louis, MO 63166; tel: (314) 444-8806; fax: (314) 444-8731 email: dicecio@stls.frb.org

Abstract

We use non-parametric distribution dynamics techniques to reassess the convergence of per capita personal income (PCPI) across U.S. states and across metropolitan and nonmetropolitan portions of states for the period 1969-2005. The long-run distribution of PCPI is bimodal for both states and metro/nonmetro portions. Furthermore, the high-income mode of the distribution across metro and nonmetro portions corresponds to the single mode of the long-run distribution across metro portions only. These results (polarization or club-convergence) are reversed when weighting by population. The long-run distributions across people are consistent with convergence. Migration and urbanization are the forces behind convergence.

JEL: O51, R11, R23.
Keywords: convergence, migration, urbanization.

1 Introduction

The study of convergence of living standards across countries and within a given country is one of the most important and fascinating issues in economics. Geographical units within a single country represent the best-case scenario for the convergence hypothesis,[1] i.e., that poorer geographical units tend to grow faster than richer ones and eventually catch up. Regions, states, departments, or prefectures share the same legal institutions, currency, and a significant part of fiscal policy; productive factors, capital and labor, can move freely. In many countries, they share the same language and cultural heritage. We analyze convergence within the U.S. along three dimensions: across states (and people therein), across metropolitan and nonmetropolitan portions of the states (and people therein), and in time.

We reassess the convergence of per capita personal income (PCPI) across U.S. states for the period 1969-2005 by using non-parametric distribution dynamics techniques. In this setting, a unimodal long-run distribution is interpretable as evidence of convergence. Notice that our approach assumes that within each state, income is distributed equally across individuals. Convergence, or lack thereof, is to be interpreted across "average individuals", one from each state. The long-run distribution of PCPI we obtain is bimodal. The emergence of polarization in the cross-section distribution of income across states and the corresponding twin-peakedness of the long-run distribution are relatively new phenomena. Quah (1996) and Johnson (2000) obtained unimodal long-run distributions using distribution dynamics techniques over 1948-1989 and 1948-1993 samples, respectively. We also consider the distribution dynamics weighted by the number of people within each state. Each state is now represented by a number of average individuals equal to its population; each individual is identical and has an income equal to the PCPI of that state. Our finding of polarization is attenuated after weighting by the population. The long-run distribution in this case is nearly single-peaked: While there is no convergence across states, there is evidence of convergence across people. Convergence across people is driven by the fact that states that are losing ground in the distribution of income also account for a declining share of the U.S. population.

To gain a deeper understanding of the evolution of income within the United States, we extend our analysis to the distribution dynamics across metropolitan (metro) and nonmetropolitan (nonmetro) portions of the states. Most metro portions start and end up with a relatively high level of income. Conversely, nonmetro portions tend to start and remain at lower levels of income per capita. Hence, the long-run distribution is twin-peaked. The high-income mode for all portions considered together corresponds to the single mode obtained by analyzing only metro portions. As in the analysis of the long-run distribution of income across states, we find that twin-peakedness disappears after weighting by the population. This occurs for two reasons. First, consistently metro and nonmetro counties account for a decreasing share of the U.S. population between 1969 and 2005: The portions of states most dissimilar account for a decreasing fraction of the

[1] There is a vast and growing literature on regional convergence within countries and across regions in Europe. See Magrini (2004) for a survey of this literature; Rey and Janikas (2005) focus their survey on the spatial elements of regional convergence. Durlauf and Quah (1999) and Durlauf, Johnson, and Temple (2005) provide general surveys on the empirics of growth and convergence.

population. Second, counties that switched from nonmetro to metro—mostly suburban areas—experienced population growth and did not lose ground relative to metro per capita incomes. We interpret this as evidence of the importance of suburbanization, as opposed to rural economic development, as the driving force behind convergence.

The analysis of long-run distributions overlooks the time-series behavior of the convergence process. In order to analyze convergence over time, we compute the distance between the cross-sectional distribution for each year and the corresponding long-run distribution. For the metro/nonmetro distributions, we find that the distances increased during the 1980s and late 1990s and declined during the 1970s and 2000s. For the distributions across states, the movements just described are dominated by an increasing distance since the early 1980s. Finally, the population-weighted distribution across states moved closer to its long-run counterpart in the early 1970s, and apart between the mid-1970s and the late 1980s. It has remained at a stable distance ever since.

The rest of the paper is organized as follows. Section 2 describes the data and methodology used. Sections 3 and 4 analyze the dynamics of the distribution of income across states and their metro and nonmetro portions, respectively. Section 5 presents our findings on the time-series behavior of the cross-sectional distributions vis-à-vis the corresponding ergodic distributions. Section 6 concludes.

2 Data and Methodology

In this section we describe our data sources and the methodology used to assess convergence.

2.1 Data

We consider convergence of PCPI across the 50 U.S. states and the District of Columbia,[2] as well as their metro and nonmetro portions over the period from 1969 to 2005. The metro portion of every state comprises the counties belonging to metropolitan statistical areas (MSAs). The nonmetro portion includes the remaining counties.[3]

As the U.S. population grows, metropolitan areas expand and engulf the surrounding counties, new metropolitan centers are formed. The Office of Management and Budget regularly reclassifies counties from nonmetro to metro. For example, in 1960 the St. Louis MSA consisted of 7 counties; by 2005 the St. Louis MSA had expanded to encompass 17 counties. In other cases new MSAs are formed when at least 50,000 people reside within one "core" city (e.g., Jefferson City, MO). All 50 states had a metro and a nonmetro portion in 1969, except for the District of Columbia, which did not have a nonmetro portion. By 2005 all of the counties in New Jersey had been classified as metro. When the classification of counties is revised, the PCPI for the metro and nonmetro portion of every state is recalculated to reflect the most recent county classification. For example, in our data the state of New Jersey does not have a nonmetro component all along our sample, even though in 1969 it had nonmetro counties.

[2] Data on PCPI for the U.S. states are available at www.bea.gov/regional/spi/.

[3] Data on PCPI for the metro and nonmetro portions of the U.S. states, as well as data at the county level, are available at http://www.bea.gov/regional/reis/. The MSA classifications are available at www.census.gov/population/www/estimates/pastmetro.html.

2.2 Methodology

We assume that the distribution of income, in logs and relative to its cross-sectional average, evolves according to the following first-order process:

$$f_{t+\tau}(y)=\int_{-\infty}^{+\infty} g_\tau(y \mid x) f_t(x)\,dx, \tag{1}$$

where f_t denotes the density[4] at time t and g_τ denotes the stochastic kernel relating the time-t and time-$(t+\tau)$ distributions. The ergodic distribution, f_∞, solves

$$f_\infty(y)=\int_{-\infty}^{+\infty} g_\tau(y \mid x) f_\infty(x)\,dx, \tag{2}$$

We estimate f_∞ non-parametrically as follows. The joint distribution $g_\tau(y, x)$ is estimated by adaptive Gaussian kernel smoothing as described in Silverman (1986, chap. 5):

$$\hat{g}_\tau(y,x)=\sum_{i=1}^{N}\frac{1/N}{h^2}\mathrm{K}\left(\frac{y-Y_i}{h},\frac{x-X_i}{h}\right), \tag{3}$$

where h is the bandwidth, K is the bivariate kernel density (the normal density in this case), and $\{Y_i, X_i\}_{i=1}^{N}$ are the observed transitions. Each transition receives an equal weight, $1/N$.[5] In the first stage we use a rule-of-thumb bandwidth (see Silverman, 1986, pp. 86-7) to obtain a pilot density estimate. In the second stage the bandwidth is inversely related to the pilot density estimated in the first stage. The conditional distribution $g_\tau(y \mid x)$ is computed dividing the joint distribution by the marginal:

$$\hat{g}_\tau(y \mid x)=\begin{cases}\hat{g}_\tau(y,x)\left[\int_{-\infty}^{+\infty}\hat{g}_\tau(y,x)\,dy\right]^{-1} & \text{if } \int_{-\infty}^{+\infty}\hat{g}_\tau(y,x)\,dy \neq 0,\\ 0 & \text{otherwise.}\end{cases} \tag{4}$$

To compute the ergodic distribution we discretize equation (2): $\varphi_\infty = \varphi_\infty \Gamma$. The stochastic matrix Γ is obtained by discretizing the stochastic kernel, g_τ. The vector φ_∞ is simply the eigenvector of Γ associated with the eigenvalue equal to one and $\hat{f}_\infty$ is recovered by undoing the discretization.[6]

In the following sections we consider as a benchmark a transition horizon of ten years, i.e., $\tau = 10$. The non-parametric estimation technique described above extends easily to weighted transitions, stochastic kernels, and ergodic distributions. In particular,

[4] We assume that the cross-sectional distribution can be described in terms of a density. See Azariadis and Stachurski (2003) or Quah (2007) for a discussion of the general case.

[5] The total number of transitions is equal to the product of the number of cross-sectional units, m, and the number of transitions per unit: $N = m \times (T-\tau)$, where T is the sample length. For example, with a 1969-2005 sample of 48 states the total number of 10-year transitions is $48 \times (37-10) = 1{,}296$.

[6] Azariadis and Stachurski (2003) provide a rigorous treatment. We followed the implementation of Johnson (2000, 2005) described in detail in a note available at: http://irving.vassar.edu/faculty/pj/finding_ergodic_density.pdf.

we introduce population weights, $\omega_i = \frac{Pop_i}{Pop_{US}}$, in the second stage[7] of the adaptive kernel smoothing by modifying equation (3) as follows

$$\hat{g}_\tau(y,x) = \sum_{i=1}^{N} \frac{\omega_i}{h^2} \mathrm{K}\left(\frac{y-Y_i}{h}, \frac{x-X_i}{h}\right).$$

In the analysis of the distribution dynamics, a unimodal ergodic distribution can be interpreted as a necessary condition for "global" convergence.[8] Conversely, the more pronounced the multi-modality of the long-run distribution is, the stronger is the evidence of polarization. A multi-modal ergodic distribution indicates that, in the long-run, there will be groups of cross-sectional units that tend to cluster at different levels of income. While there will be convergence within each group, there will be no convergence across different groups and, thus, there will be no global convergence.

To assess the evolution of the cross-sectional distribution of income we construct time series of the distance between the cross-sectional distribution at time t and the corresponding ergodic distribution. We estimate the cross-sectional distribution for every year, $\hat{f}_t$, by univariate kernel smoothing.[9] Finally, we compute the L_1 distance from the estimated long-run distribution, $\hat{f}_\infty$, and time-t distribution, $\hat{f}_t$, as follows:

$$d_t = \int_{-\infty}^{+\infty} \left| \hat{f}_t(x) - \hat{f}_\infty(x) \right| dx, \ t = 1969, \ldots, 2005. \qquad (5)$$

3 Convergence Across States

The analysis of regional convergence within the United States is a classical empirical exercise undertaken by many authors. In this section we reassess the common wisdom that there is convergence of income across states. In the literature on per capita income convergence for U.S. states two primary methods have been employed.[10] The most popular, although not necessarily preferred method, is the regression approach: Starting with the work of Barro and Sala-I-Martin (1991, 1992) authors have generally agreed there has been convergence between U.S. states (e.g., Holtz-Eakin, 1993; Garofalo and Yamarik, 2002). Similarly, research employing time-series methods has found evidence of convergence between 1929 and 1990 (e.g., Carlino and Mills, 1993, 1996). Despite this consensus, Bernard and Durlauf (1996) note that neither methodology is likely to yield unambiguous conclusions about convergence: Each approach employs a different meaning of convergence and makes different assumptions about the properties of the data. Bernard and Durlauf show that cross-section tests can reject the null hypothesis of

[7] It is possible to introduce weighting in the first stage of the estimation as well. However, this produces very narrow bandwidths and undersmoothing.

[8] If there is low intra-distribution mobility, the cross-sectional units in the left (right) tail of the distribution will remain below (above) the unique mode. A long-run distribution where rich remain rich and poor remain poor is not consistent with convergence, even if most of the cross-sectional units have similar levels of income.

[9] We adopt an adaptive Gaussian kernel smoother, weighting by the population in the second stage of the estimation when appropriate.

[10] See Magrini (2004) for a complete summary of the literature on regional convergence.

no convergence even for data generated by economies with different long-run steady states, and time-series tests may be invalid if the data are driven by transition dynamics. Quah (1993) notes, the main problem that has plagued the regression approach to convergence is that it gives little if not any information on the intra-distribution dynamics. Additionally, Rey and Janikas (2005) and Rey and Dev (2006) recognize that spatial effects can bias cross-sectional convergence measures. In order to resolve these problems authors have employed variants of the methodology laid out in this paper. Quah (1996) and Johnson (2000) have applied this methodology for convergence across states and, as noted earlier, both of these papers present unimodal ergodic distributions, implying convergence. More recently, Webber, White, and Allen (2005) use measures of concordance and find the shape of the (actual) distribution suggests that the states are converging.[11]

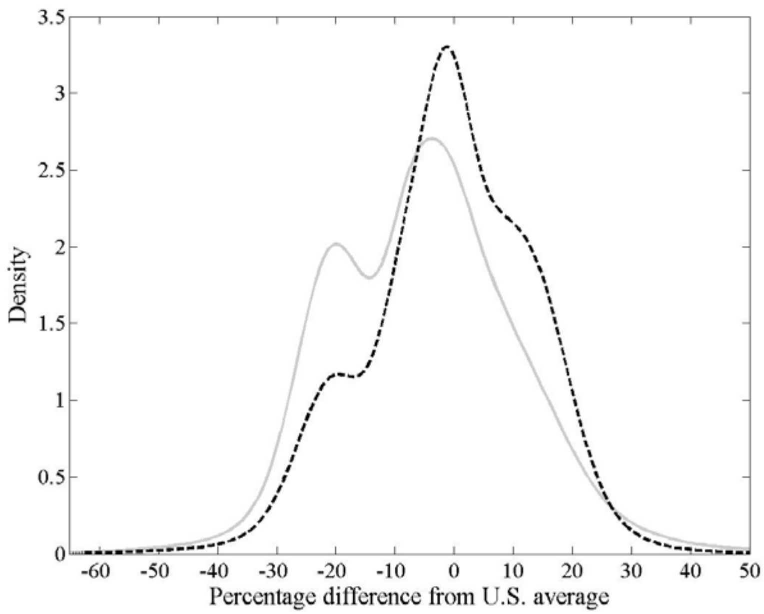


Figure 1: Ergodic distribution of PCPI across states: unweighted (solid gray) and population weighted (dashed black).

In contrast with the existing literature we find a bimodal long-run distribution. However, when we weight the transitions by the corresponding population single-peakdness is restored. Figure 1 portrays the ergodic distributions of PCPI across states, unweighted and weighted by population. The shapes of the ergodic distributions portrayed are robust to the perturbations described in Table 1. Notice that the unweighted distribution is bimodal. The lowest mode corresponds to a PCPI 20.1 percent below the U.S. average. The highest mode is associated with a PCPI 3.8 percent below the cross-sectional average. The bimodal ergodic distribution in Figure 1 is in sharp contrast with the unimodal long-run distribution obtained by Johnson (2000). Two main factors are responsible for this difference in the number of modes. First, Johnson's sample is from

[11] Some pairs of states were identified as exhibiting divergence. Furthermore, Webber, White, and Allen (2005) present mixed evidence leaning toward divergence in the later portion of their sample.

1948 to 1993, while ours covers the period 1969 to 2005. In particular, over the past decade the cross-sectional distribution of PCPI has become increasingly bimodal, which is reflected in the shape of the ergodic distribution. The fact that we do not include the period from 1948 to 1968 in our benchmark analysis does not affect, qualitatively, our results. Second, Johnson (2000) considers only three 15-year transitions while we consider all 27 10-year transitions in our sample.

	Geographical unit	Sample	Kernel	Transformation	Transition Period
Baseline	50 states (plus DC)	1969-2005	Gaussian	logs (scaled by U.S. PCPI)	$\tau = 10$
Alternative	48 states	1948-2005	Epanechnikov	Levels (scaled by U.S. PCPI)	$\tau = 5, 15$

Table 1: Robustness analysis

Following Quah (2007), we also consider the distribution dynamics weighted by the number of people within each state. The population-weighted ergodic distribution can be interpreted as the long-run distribution across people, as opposed to states. Notice that the implicit assumption is that all residents within each state have the same level of PCPI. The ergodic distribution of PCPI across people is still bimodal, but the low-income mode is much less pronounced. The standard deviation of the population-weighted ergodic distribution is 11 percent lower than the standard deviation of the unweighted long-run distribution. The average of the population-weighted distribution is closer to the U.S. average income than the average of the unweighted distribution is.

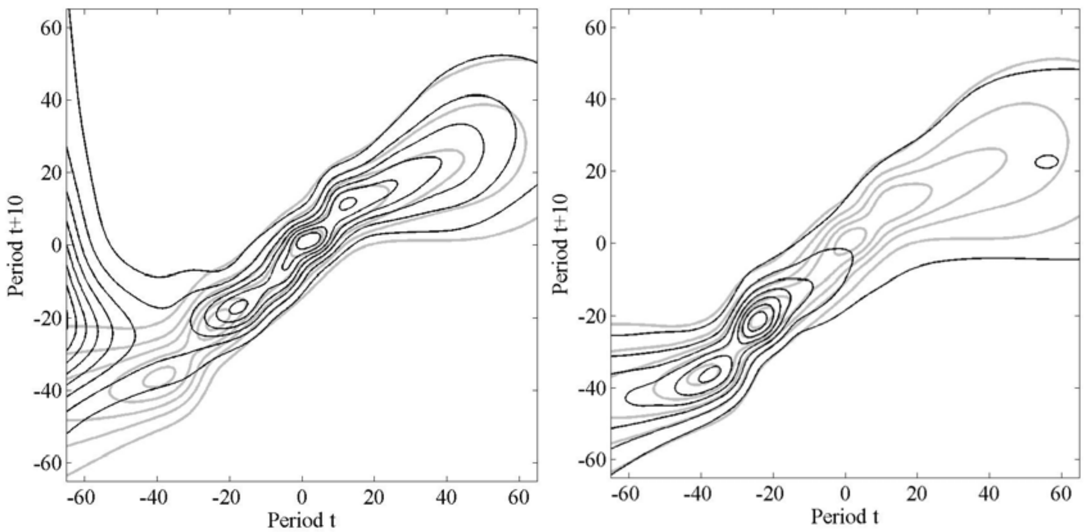
Convergence across people is driven by the fact that states that are losing positions in the income distribution are also losing population. For example, Ohio in 1969 had the 15th highest income at 8 percent above the national average. By 2005 Ohio lost ground: It occupied the 30th place with an income of 4.5 percent below the national average. At the same time, Ohio's population declined from 5.35 percent of the total U.S. population in 1969 to below 4 percent in 2005. Conversely, states climbing the ladder of the income distribution were gaining population, contributing to convergence. Colorado was the 22nd state in terms of PCPI in 1969 and climbed to the 9th place by 2005. During the same period, Colorado's population share increased from 1.1 to 1.6 percent.[12]

4 Convergence Across Metro and Nonmetro Portions of States

To provide further insight into the dynamics of income within the United States, this section considers the distribution dynamics at a finer level of geographical aggregation,

[12] In general the Sun Belt states (e.g., CA, NM, and AZ) experienced relative population and income gains, while the Rust Belt states (e.g., OH, PA, IN, and MI) experienced relative declines in income and population.

i.e., the metro and nonmetro portions of each state. The common wisdom is that there is convergence within metro and nonmetro regions, but not between them (e.g. Nissan and Carter, 1999, 2005; Hammond, 2004, 2006). Figure 2a portrays the contours of the stochastic kernel[13] for the metro portions of states (black lines), together with contours of the stochastic kernel for all portions of states (gray lines). Figure 2b contrasts the stochastic kernel for all portions (gray) and the one for nonmetro portions only (black). All three stochastic kernels have most of the mass along the 45º line, implying high persistence. Most of the transition for nonmetro portions start and end up with a relatively low level of income, i.e., the mass is concentrated on the bottom-left corner in Figure 2b. Vice versa, metro portions tend to start and remain at levels of income higher than the U.S. average.



(a) Metro portions, black; all portions, gray. (b) Nonmetro portions, black; all portions, gray.

Figure 2: Contour plots of stochastic kernels.

The ergodic distributions associated with the stochastic kernels discussed above are portrayed in Figure 3. When metro and nonmetro portions are considered together, the long-run distribution is bimodal. The two modes correspond to PCPI values of 17.1 and 0.3 percent below the U.S. PCPI, respectively. For metro portions only, the ergodic distribution is single-peaked with a maximum at 1.3 percent above the U.S. PCPI. On the other hand, the long-run distribution of PCPI across nonmetro portions of states is twin-peaked. The two peaks are located at values of income 21.2 and 34.9 percent below the national average.[14]

[13] The contour lines in Figure 2 are curves along which the corresponding stochastic kernel has a constant value. A section of the stochastic kernel along the "Period t+10" axis at a given "Period t" value is the distribution of PCPI at time t +10, conditional on the time-t value.

[14] We checked the robustness of the results discussed above against the perturbations in Table 1. Data availability does not allow us to consider the extended sample 1948-2005 for metro/nonmetro regions.

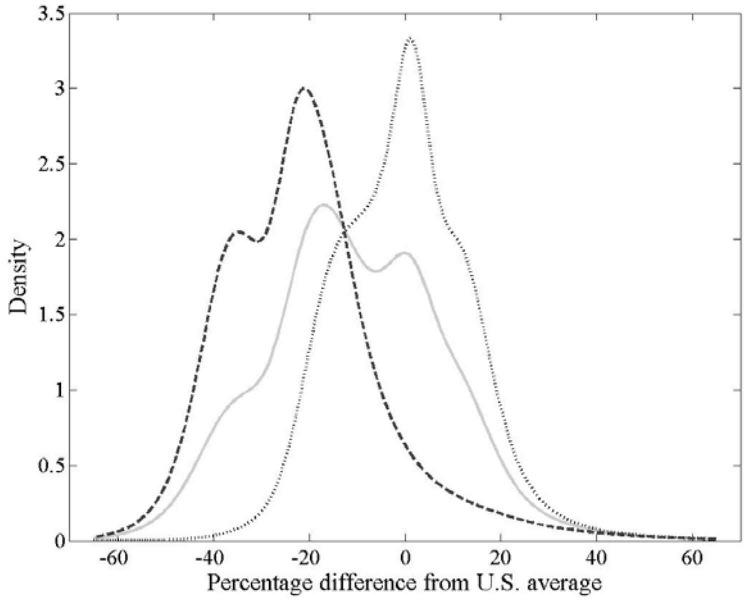


Figure 3: Ergodic distribution of PCPI across metro and nonmetro portions of states: all portions together (light gray solid), only nonmetro portions (dashed dark gray), and only metro portions (dotted black).

Figure 4 contrasts the ergodic distribution across metro and nonmetro regions with the long-run distribution across people in those regions. The latter distribution is unimodal and has a 9 percent lower standard deviation. The mode is very close to the nation-wide average PCPI. This long-run behavior is likely a result of population movements, which we explore further.

Using county-level data, we are able to identify the counties that experienced significant population growth over the sample, specifically those counties that were reclassified from nonmetro to metro (henceforth, transition counties). In 1973 the average PCPI of these transition counties reached a high point of 85 percent of metropolitan areas; by 2005 these counties' average income declined slightly to 83 percent, relative to those counties that were metropolitan over the entire sample. Conversely, in 1973 the counties that remained nonmetro over the entire sample had an average PCPI of 87 percent of that of metro areas. By 2005 the areas that remained nonmetro had seen their average PCPI decline to 74 percent of that of the metro areas. These two different experiences suggest that urbanization has slowed the rate of divergence.

The easiest way to capture the effect of urbanization is by holding the land area of the MSAs fixed at their 1969 values (core MSAs), calculating the percentage of the population living within those areas, and then fixing the land area of MSAs at their expanded 2005 levels and calculate the population shares (see Figure 5). Between 1969 and 2005, the share of population living in core MSAs declined from 63 percent to around 60 percent, denoted by the dashed line. Over this time period, transition counties experienced strong population growth. Population in transition counties grew three times as fast as in nonmetro counties, twice as fast as in core metro counties, and 1.7 times as fast as overall U.S. population. Observing the 2005 dimensions of the MSAs, we see a steady increase in the share of the population living in urban areas.

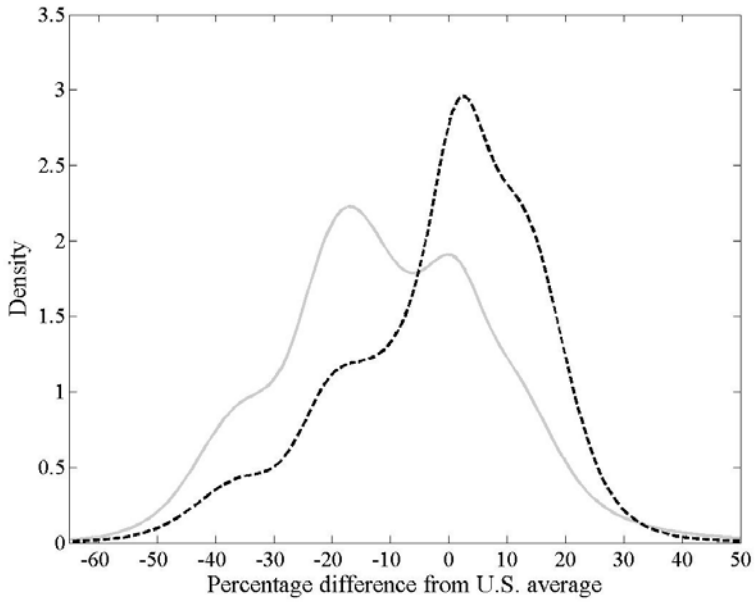


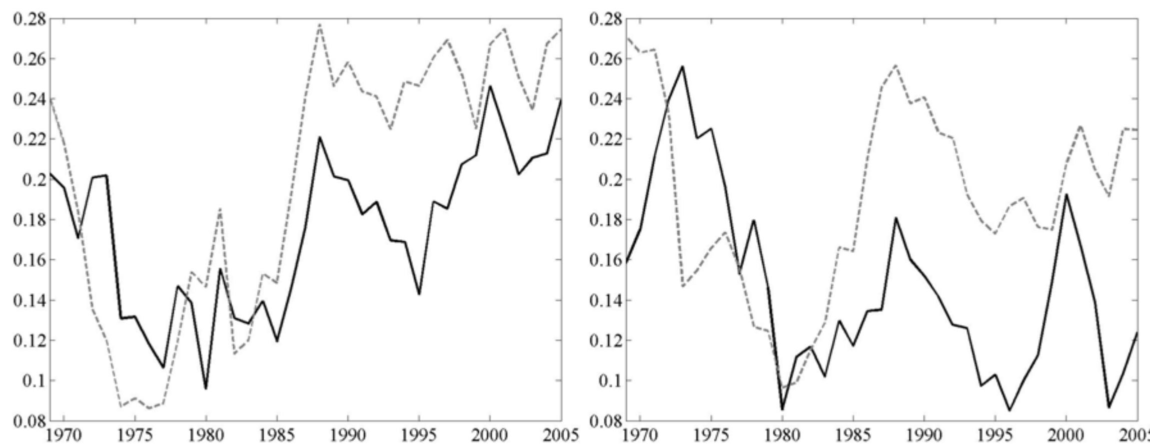
Figure 4: Ergodic distribution of PCPI across metro and nonmetro portions of states: unweighted (solid gray) and population weighted (dashed black).

Figure 5: Share of U.S. population living in MSAs: 2005 classification (black solid) and 1969 classification (gray dashed).

5 Convergence in Time

The analysis of long-run distributions overlooks the time-series behavior of the convergence process. Several authors have argued that convergence, or lack thereof, is not a smooth process in time. With respect to convergence across states, Coughlin and Mandelbaum (1988) find that the convergence trend since the 1920s is reversed in the late 1970s. Webber, White, and Allen (2005) find convergence across states from 1930 until 1975, and limited evidence of divergence from 1975 through 2000. For metro and nonmetro regions, Nissan and Carter (1999) identify the 1970s and the 1990s as periods of convergence. On the other hand, Nissan and Carter (2005) determine that convergence occurred only in the early portion of the 1970s and there was "significant divergence [between metro and nonmetro regions] thereafter."

In this section we describe the behavior of the distances between the cross-sectional distributions of PCPI in every year and the long-run distributions estimated in the previous sections. All distances are computed using equation (5).



(a) States. (b) Metro/nonmetro portions.

Figure 6: Distance of the cross-sectional distribution from the ergodic distribution: unweighted (black, solid lines) and population weighted (gray, dashed lines).

Figure 6a portrays distances of the distributions of PCPI across states from their long-run counterparts discussed in Section 3 above. The solid black line is the distance for the unweighted distributions; the dashed gray line is the distance for the population-weighted distributions. The distribution across states moves closer to its long-run counterpart in the late 1970s. Since the early 1980s the cross-sectional and the long-run distribution have moved apart. This general diverging trend was mitigated by episodes of convergence in the early 1990s and early 2000s. The distribution across people in states approached its long-run counterpart in the early 1970s, diverged from it between the mid-1970s and the late 1980s, but it has remained at a roughly stable distance ever since.

Figure 6b shows that the evolution over time of the cross-sectional distributions across metro/nonmetro portions of states and across people in metro/nonmetro areas: Both cross-sectional distributions moved closer to their long-run distributions in the late 1970s, the early 1990s, and the early 2000s.

Notice that the trends outlined in Figures 6a and 6b do not necessarily correspond to the periods of convergence/divergence within the cross-sectional distributions analyzed by Coughlin and Mandelbaum (1988) and Webber, White, and Allen (2005) for states and by Nissan and Carter (1999, 2005) for metro and nonmetro regions. For example, a declining distance over time does not imply convergence across states or metro/nonmetro regions. Instead, it simply reflects that the cross-sectional distribution is becoming closer to the ergodic distribution. When the ergodic distribution is multi-modal, this could be associated with states or metro/nonmetro regions diverging from each other.

6 Conclusions

We have examined the convergence of per capita personal income across U.S. states and across metro and nonmetro portions of states. A common theme is that, while there is no convergence across geographical units at either level, there is convergence across people.

Contrary to previous results of convergence across states, our finding is of a bimodal long-run distribution. When population-weighting is introduced, we obtain a nearly unimodal ergodic distribution. In other words, convergence across people is driven by the fact that states climbing ranks in the distribution of income are also attracting a larger share of the population.

Regarding the metro and nonmetro portions of states, there is a strong case for club-convergence. Metro portions of states will converge to incomes near the national average, while nonmetro portions will converge to lower incomes. The long-run distribution across people, however, is consistent with convergence. The driving force behind convergence across people is urbanization. Core metro portions of the states tend to grow faster than the national average, but their population share is declining. Conversely, consistently nonmetro portions experience income and population declines. The portions of the states that gained in population share are those that became parts of MSAs over our sample, typically as a result of suburbanization. Also, the average income of these transitioning portions of states remained closer to the national average.

Our results provide evidence that the idea of preserving rural economies while achieving significant gains in per capita income (or slowing divergence) in the long run appears to be far-fetched. These results also support the ideas of Gottmann (1961) and Lang and Dhavale (2005) that cities that are expanding, running into small towns, and merging into "super cities" will form the engine of economic growth in the future.

References

Azariadis, C., and J. Stachurski (2003): "A Forward Projection of the Cross-Country Income Distribution," Discussion Paper No. 570, Institute of Economic Research, Kyoto University.

Barro, R. J., and X. Sala-I-Martin (1991): "Convergence across States and Regions," *Brookings Papers on Economic Activity*, 0(1), 107-58.

Barro, R. J., and X. Sala-I-Martin (1992): "Convergence," *Journal of Political Economy*, 100(2), 223-51.

Bernard, A. B., and S. N. Durlauf (1996): "Interpreting Tests of the Convergence Hypothesis," *Journal of Econometrics*, 71(1-2), 161-73.

Carlino, G. A., and L. O. Mills (1993): "Are U.S. Regional Incomes Converging?: A Time Series Analysis," *Journal of Monetary Economics*, 32(2), 335-46.

Carlino, G. A., and L. O. Mills (1996): "Convergence and the U.S. States: A Time-Series Analysis," *Journal of Regional Science*, 36(4), 597-616.

Coughlin, C. C., and T. B. Mandelbaum (1988): "Why Have State Per Capita Incomes Diverged Recently?," Federal Reserve Bank of St. Louis *Review*, 70(5), 24-36.

Durlauf, S. N., P. A. Johnson, and J. R. Temple (2005): "Growth Econometrics," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf, vol. 1A, pp. 555-677. Elsevier Science, North-Holland, Amsterdam; New York and Oxford.

Durlauf, S. N., and D. T. Quah (1999): "The New Empirics of Economic Growth," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1A, pp. 235.308. Elsevier Science, North-Holland, Amsterdam; New York and Oxford.

Garofalo, G. A., and S. Yamarik (2002): "Regional Convergence: Evidence from a New State-by-State Capital Stock Series," *Review of Economics and Statistics*, 84(2), 316-23.

Gottmann, J. (1961): *Megalopolis: The Urbanized Northeastern Seaboard of the United States.* Twentieth-Century Fund.

Hammond, G. W. (2004): "Metropolitan/Non-Metropolitan Divergence: A Spatial Markov Chain Approach," *Papers in Regional Science*, 83(3), 543-63.

Hammond, G. W. (2006): "A Time Series Analysis of U.S. Metropolitan and Non-Metropolitan Income Divergence," *Annals of Regional Science*, 40(1), 81-94.

Holtz-Eakin, D. (1993): "Solow and the States: Capital Accumulation, Productivity, and Economic Growth," *National Tax Journal*, 46(4), 425-39.

Johnson, P. A. (2000): "A Nonparametric Analysis of Income Convergence Across the US States," *Economics Letters*, 69(2), 219-23.

Johnson, P. A. (2005): "A Continuous State Space Approach to *Convergence by Parts*," *Economics Letters*, 86(3), 317-321.

Lang, R., and D. Dhavale (2005): "America's Megapolitan Areas," *Land Lines*, 17(3), 1-4.

Magrini, S. (2004): “Regional (di)convergence,” in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J.-F. Thisse, pp. 2741-96. Elsevier Science, North-Holland, Amsterdam, New York and Oxford.

Nissan, E., and G. Carter (1999): “Spatial and Temporal Metropolitan and Nonmetropolitan Trends in Income Inequality,” *Growth and Change*, 30, 407-29.

Nissan, E., and G. Carter (2005): “Decomposition of Regional Metropolitan and Non-metropolitan Income Inequality,” *Journal of Economics and Finance*, 29(1), 73-84.

Quah, D. T. (1993): “Galton’s Fallacy and Tests of the Convergence Hypothesis,” *Scandinavian Journal of Economics*, 95(4), 427-43.

Quah, D. T. (1996): “Empirics for Economic Growth and Convergence,” European Economic Review, 40(6), 1353-75.

Quah, D. T. (2007): *Growth and Distribution*. unpublished manuscript, London School of Economics.

Rey, S. J., and B. Dev (2006): “σ-Convergence in the Presence of Spatial Effects,” *Papers in Regional Science*, 85(2), 217-34.

Rey, S. J., and M. V. Janikas (2005): “Regional Convergence, Inequality, and Space,” *Journal of Economic Growth*, 5(2), 155-76.

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

Webber, D. J., P. White, and D. O. Allen (2005): “Income Convergence Across U.S. States: An Analysis Using Measures of Concordance and Discordance,” *Journal of Regional Science*, 45(3), 565-89.